WITHIN–PERSON VARIATION IN CORONARY RISK FACTORS: IMPLICATIONS FOR THE AETIOLOGY AND PREVENTION OF CORONARY HEART DISEASE

THESIS
presented for the degree of DOCTOR OF PHILOSOPHY
in the Faculty of Medicine
(Field of Study - Epidemiology).

by Jonathan R Emberson

Department of Primary Care and Population Sciences,
Royal Free and University College Medical School,
University of London

August 2004

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Epidemiological studies clearly demonstrate the importance of numerous risk factors for coronary heart disease (CHD), including blood lipids, blood pressure, cigarette smoking and physical inactivity. These factors are widely believed to account for only around 50% of CHD cases. However, "within–person" variation in coronary risk factors can affect the size and even direction of estimated aetiological relationships, and though these effects have been explored for the univariate relations of blood pressure and blood cholesterol, much uncertainty remains. In this thesis, data from the British Regional Heart Study, a prospective study of cardiovascular disease in middle–aged British men, is used to investigate the extent and effects of "within–person variation" in a range of coronary risk factors. The effects on estimated relations with CHD are examined and the combined importance of the major risk factors to CHD risk assessed. The potential effectiveness of different CHD prevention strategies, and the size and cause of social inequalities in CHD are also estimated. The findings reveal a high degree of within–person variation in both established and novel coronary risk factors. Taking within–person variation into account, CHD risk–relations for blood lipids, blood pressure, cigarette smoking and physical inactivity increase in magnitude; though the estimated protective effect from moderate alcohol intake is reduced. After correction for within–person variation, blood cholesterol, blood pressure and cigarette smoking together account for at least 75–80% of CHD cases in British men. Moderate population–wide improvements in these risk factors could therefore greatly reduce population levels of CHD, while "high–risk" strategies, unless applied to a large proportion of the population, are likely to have only a limited effect. Narrowing social inequalities in CHD would also have a comparatively modest effect on CHD compared with population–wide control of the key causal coronary risk factors.

DECLARATION OF AUTHORSHIP

I have developed the research hypotheses that are addressed in this thesis, and all the analysis and commentary that is presented is my own work. I have used data from the British Regional Heart Study (an ongoing cardiovascular study that was initiated in 1978). I played no role in the design or the data collection for this study.

Candidate: Jonathan Robert Emberson

Signed:

Date: 24 - 02 - 05

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to Dr Richard Morris and Professor Peter Whincup for their continued insight, advice, support and encouragement on all aspects of this work. I would also like to thank Gerry Shaper, Mary Walker and Shah Ebrahim for their valuable comments on various aspects of this work, and Andy Thomson for providing numerous data files on the British Regional Heart Study. In addition, I would like to thank Ellen Bleasdale for all the support she has given me during the past three years, and also for proof reading this thesis.

The British Regional Heart Study is funded by the British Heart Foundation and receives additional support from the Department of Health (England). I would like to thank the British Heart Foundation for awarding me a Junior Research Fellowship, without which this work would not have been possible.

LIST OF ABBREVIATIONS

Epidemiological studies

BRHS	British Regional Heart Study

Disease outcomes

CHD	Coronary heart disease
CVD	Cardiovascular disease
MI	Myocardial infarction

Risk factors and measurements

BMI	Body mass index (kg/m^2)
DBP	Diastolic blood pressure (mmHg)
ECG	Electrocardiogram
FEV1	Forced expiratory volume in one second (l)
HDL	High density lipoprotein cholesterol (mmol/l)
LDL	Low density lipoprotein cholesterol (mmol/l)
SBP	Systolic blood pressure (mmHg)
TC	Serum total cholesterol (mmol/l)

Statistical terms

CI	Confidence interval
HR	Hazard ratio
IQR	Interquartile range
PARF	Population attributable risk fraction
RDR	Regression dilution ratio
RR	Relative risk
SD	Standard deviation

Contents

List of chapter subsections

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Summary

Despite impressive falls over the latter half of the 20th century, coronary heart disease (CHD) remains the single most common cause of death in the United Kingdom and is an increasingly important problem worldwide. Over the last 50 years, epidemiological studies have clearly demonstrated the importance of numerous risk factors for CHD, such as adverse blood lipids, high blood pressure, and "lifestyle characteristics" such as cigarette smoking and physical inactivity. Many other "novel" risk factors have also subsequently been implicated as potentially important. However, the contributions of these established and novel risk factors to CHD risk are not completely understood. In particular, it is unclear what contribution the established major coronary risk factors make to individual variation in CHD risk. The effects that "within–person variation" in coronary risk factors have on estimated disease relationships (often resulting in underestimation of true relationships) has added to this uncertainty, and though these issues have been explored for the most established risk factors such as blood pressure and blood cholesterol, they have usually only been considered in a univariate setting. In this thesis, data from the British Regional Heart Study (BRHS), a prospective study of cardiovascular disease established in 1978–80 and comprising 7,735 middle–aged British men followed for cardiovascular morbidity and all–cause mortality for over 20 years, are used to investigate the extent and effects of within–person variation in a range of known coronary risk factors. In particular, the effect that within–person variation has on: (1) estimated relationships with CHD; (2) the combined importance of the three strongest risk factors to population levels of CHD;

(3) the potential effectiveness of different CHD prevention strategies; and (4) the true size and explanation of social inequalities in CHD, are examined. Particular strengths of the BRHS that are important for addressing these research hypotheses are the socially and geographically representative nature of the study population, the non–interventional nature of the study, the high rates of follow-up, and, crucially, the availability of repeated measurements of both physical and biochemical risk factors and lifestyle characteristics at various points throughout the study period.

1.2 Background

1.2.1 The coronary heart disease epidemic

Diseases of the heart and circulatory system, and in particular those diseases falling under the definition of coronary heart disease (CHD), are by far the most common cause of death in Western societies and are an increasingly important cause of death worldwide.[1;2] In 2001, coronary (ischaemic) heart disease accounted for an estimated 13% of all deaths worldwide (representing approximately 3.8 million men and 3.4 million women), around the same number as all cancers combined, nearly twice the number of deaths due to respiratory infections and nearly three times the number of HIV/AIDS deaths.[1] In the United Kingdom, coronary heart disease accounts for around one half of all circulatory deaths (three fifths in men) and is the cause of death for about one in four men and one in six women; over a third of these occur in individuals under 75 years of age.[3] However, at the beginning of the 20th century, CHD was only the fourth most common cause of death in the United Kingdom behind pneumonia, tuberculosis and diarrhoeal disease. By 1910, it had already reached first place with substantial increases in the proportion of deaths attributable to CHD being observed throughout the first half of the 20th century (see Figure 1.1). With the exception of a drop in CHD mortality during the 1940's (thought to be caused by the effects of rationing during the Second World War), CHD continued to increase in the UK throughout the 20th century until the 1970's, when the epidemic reached its peak. The last 30 years have seen a substantial fall in the incidence of CHD in the United Kingdom, however rates in the UK have not fallen as dramatically as those in other countries (see Figure 1.2). In addition, the absolute number of people who die from CHD in the UK (and many other Western societies) has changed very little despite

these falls in incidence. Furthermore, the number of chronically ill CHD patients may even be increasing in these countries as the population ages and survival improves.[4;5] For this reason, the overall burden of CHD, and other vascular diseases, in developed countries is unlikely to decrease in the coming years, and may even increase.[6]

In developing countries, in particular India and those in South East Asia, a new epidemic of CHD has been emerging,[7–9] and while the relative contribution of CHD to all mortality in these countries is currently lower than that for developed countries, the total CHD mortality from these countries contributes a substantially greater share to the global burden of CHD than that from developed countries (because of the much larger populations in these countries). In 1998, for instance, it was estimated that over 85% of the global burden of cardiovascular disease (predominantly CHD) was from low and middle–income countries.[7] Furthermore, the number of deaths from CHD in these countries is predicted to increase substantially in the coming decades, both with decreasing risks of death from other causes (particularly infections / nutritional disorders) and increasing levels of adverse CHD risk factors. In fact, by 2020, it is estimated that the number of deaths from CHD in developing countries will be over twice that observed in 1990 (compared with roughly a 40% expected increase in developed countries).[7]

1.2.2 Pathophysiology

Coronary heart disease covers a range of clinical syndromes that include chronic conditions such as angina pectoris, and acute events such as myocardial infarction, myocardial insufficiency, and sudden death, but regardless of the clinical manifestations of CHD, the common underlying pathology is atherosclerosis of the coronary arteries resulting in ischaemia of the myocardium.[10] Arteries are made up of an inner lining called the endothelium (an elastic membrane that allows the artery to expand and contract), a layer of smooth muscle, and a layer of connective tissue. Atherosclerosis is a disease of the endothelium in which the arterial channel (lumen) is narrowed by the formation of raised patches of atheromous plaque that develop in the endothelium. These plaques (also known as lesions or atheroma) are built up over a period of time and consist of a mixture of low-density lipoproteins, fibrous tissue, decaying muscle cells, blood platelets, calcium and cholesterol that are deposited in the walls of coronary arteries. As plaques grow in size, the inner layer of the artery wall thickens, the artery narrows and the amount of blood that can

flow through the artery is reduced. This reduces the amount of oxygen reaching the heart causing myocardial ischaemia and can often lead to symptoms such as exertional chest pain (angina pectoris). However, the real danger of plaque lies in their tendency to fissure (rupture) and ulcerate.[11] When this happens the lipid content of the plaque mixes freely with the blood, resulting in the formation of a thrombus (blood clot), which may occlude the artery, leading to an acute "major" coronary event (myocardial infarction or sudden death). Even if such an occlusion does not occur however, fissure healing may result in a larger plaque and a progression of arterial obstruction. A variety of factors accelerate the development of coronary atherosclerosis, including increased blood cholesterol, increased blood pressure, cigarette smoking and diabetes[12] (see chapter 2).

1.2.3 Clinical manifestations of CHD

Coronary heart disease may present as an acute or severe event (including myocardial infarction and sudden death originating from the myocardium); as a chronic event (angina); or as a range of other conditions including, for example, heart failure and atrial fibrillation. For acute and chronic myocardial ischaemia, the first symptom is usually chest pain, though the disease may be well advanced by the time these symptoms are manifest. In fact, symptomatic CHD is usually thought of as the end–process of a disease pre–dated by long–standing sub–clinical atherosclerosis.[13;14] As previously described, a myocardial infarction occurs when the blood supply to the heart is blocked by a blood clot in one of the coronary arteries, leading to damage or death (infarction) of the affected tissues. This will usually result in severe chest pain behind the sternum (breast bone), often radiating towards the left arm, and can quickly result in death depending on the severity of the episode and the speed with which treatment is received. Angina (shortened from "angina pectoris" from the Latin for "tight chest") is also caused by a restriction, or slowing down, of the blood flow to the heart, usually due to the obstruction of a coronary artery by atheroma. The symptoms of angina are similar to those of a myocardial infarction (such as a constricting feeling in the centre of the chest), and are caused when the narrowed artery fails to supply enough blood to the heart, usually when demand is increased, such as during periods of exertion. The pain from angina usually recedes within a few minutes of ceasing exertion.

1.3 Epidemiological studies and CHD

1.3.1 Risk factors (and causal factors)

A risk factor for coronary heart disease may be defined as any characteristic found in healthy individuals that is related to the subsequent development of CHD. In general terms, risk factors may be grouped into one of the following categories:

1. Factors that are causally related to disease risk through known biological pathways, either by directly causing a change in the risk of disease (such as the influence of cigarette smoking on lung cancer risk), or else leading to a change in disease risk because of their effects on other risk factors (for instance, the influence of body mass index on disease risk may be partially attributable to its effects on blood pressure). Removing or reducing exposure to causally related risk factors should lead to a subsequent reduction in disease risk, though the extent of "reversibility" will usually depend on the duration of exposure to the risk factor;

2. Factors that act as "surrogate markers" for one or more other influences that cause disease, but do not themselves cause disease (for instance, income may be found to be related to the risk of a particular disease occurring, not because it causes disease, but because it is related to other factors that do, such as diet, social factors and the propensity to smoke cigarettes). When the relationship between these factors and disease can be accounted for by its relationships with these other "confounding" factors, we say that the risk factor is not "independently" related to disease risk.

The key feature that distinguishes type 1 from type 2 above is that in the latter case, the likelihood of disease may not necessarily be reversed by altering exposure to the risk factor (indeed the risk factor may not even be reversible, e.g. gender). Assessing the extent to which a relationship between a risk factor for CHD is causal and the extent to which it may be due to "confounding" by other factors has long been one of the aims of the observational epidemiological study. Observational epidemiological studies provide *the* key source of information for estimating the magnitude of risk–relationships, as they allow assessment of differences in disease risk corresponding to long–term differences in risk factor exposure (whereas trials, for instance, may not identify true risk associations because of, for instance, problems with compliance and incomplete reversibility). However, it is usually only by

combining evidence from pathological studies, observational epidemiological studies and randomised controlled trials (to assess the extent that risk can be reduced through removal of the risk factor), that a risk factor's "causality" may truly be established beyond doubt.

1.3.2 Risk associations and "within–person variation"

One of the primary objectives of the observational study is to examine how exposure to one or more characteristics (risk factors) influences the risk of developing disease. In the prospective cohort study this may be done by studying individuals over a long period of time and relating their personal characteristics prospectively to the occurrence of particular disease "events". A variety of statistical measures are then available with which one may quantify the nature and strength of these relationships (for example, the relative risk; see chapter 3). The magnitude of these estimates, as well as the precision with which they can be estimated, greatly influences the relative importance that is likely to be given to control of the risk factor in the population and may also influence whether other, perhaps as yet unknown, risk factors are thought to play a critical role.

In order to estimate these associations, prospective epidemiological studies have historically used single "baseline" assessments of risk factor exposure and related these to subsequent disease events. However, for single continuous risk factors that display linear "dose–response" relationships with disease risk, it is now well recognised that the use of baseline measures in analyses tends to result in underestimation of their importance.[15–18] This is because differences in risk factor levels observed among a group of individuals at a "baseline" examination tend to be larger than the true underlying differences that exist among the individuals over a period of time (because of "within–person variation": measurement errors; short term deviations from average levels; and risk factor changes over time). Consequently any observed differences in disease risk that are displayed by baseline levels of the risk factor are likely to underestimate the true differences in risk corresponding to that range of exposure levels. This becomes particularly important when the risk factor is causally related to the risk of disease because it affects the relative weight that is applied to controlling the risk factor in the population. The tendency to underestimate associations between risk factors and disease when basing analyses on single risk factor measurements has become known as "regression dilution bias",[17;19;20] however the wider implications of "within–person variation" in risk factors are not limited only to continuous

risk exposures. Risk factors that are categorical in their nature (such as cigarette smoking) are also subject to measurement errors and changes over time, and only by assessing and accounting for these effects can their true importance in an epidemiological study be evaluated. However, with the exception of studies of the effects of misclassification in dichotomous factors on risk associations,[21–24] relatively few studies have been able to examine these influences. A further complication arises when analyses are adjusted for more than one risk factor subject to within–person variation. Under such conditions, the practice of applying univariate correction methods may not be appropriate, as the extent and direction of any bias depends on the correlation structure between the variables and their individual relationships with the disease.[25–27] These issues give rise to the following important questions:

(Q1) What is the extent of within–person variation in continuous and categorical risk factors for coronary heart disease?

(Q2) What effect does this variation have on estimated relationships with coronary heart disease risk?

1.4 Studying the effects of within–person variation

1.4.1 Impact on understanding the causes of CHD

Epidemiological studies have overwhelmingly demonstrated the importance of environmental and behavioural conditions in the development of atherosclerosis and CHD, and in particular the role of three risk factors: serum total cholesterol, blood pressure and cigarette smoking.[28;29] The contribution of these, and other, established risk factors to CHD risk has been debated however, with many authors claiming that only approximately half of CHD cases may be "explained" by adverse levels of these factors.[30–40] Numerous novel risk factors have been proposed as potentially important contributors to CHD risk, with varying degrees of evidence to support them (see Chapter 2). However the effect that within–person variation in the established risk factors has on their estimated contribution to overall CHD risk has not been adequately assessed. Though several studies have

suggested that the true contribution of the established risk factors to CHD risk may be somewhat greater than one half,[41-47] these have not taken within–person variation into account. There is therefore a need to address the following issue:

> (Q3) How does within–person variation in the established coronary risk factors affect our understanding of the relative contribution of different causes of CHD?

1.4.2 Impact on CHD primary prevention strategies

Two approaches to the primary prevention of CHD are widely recognized - the "high–risk" and the "population" approaches.[48] Historically, high–risk approaches focussed on the control of individual risk factors for CHD, including cigarette smoking, high blood pressure and high blood cholesterol level.[48;49] However, the potential effectiveness of high–risk strategies has increased markedly during the past two decades, both with the availability of scoring systems to detect absolute CHD risk,[50] and with the advent of several treatments which produce marked reductions in CHD risk in high risk subjects.[51] In contrast, population strategies to prevention focus on tackling the determinants of disease by seeking to cause a downwards shift in the population distributions of the strongest risk factors. The great strength of the population strategy lies in the observation that (in Western countries) most CHD cases occur not amongst the small number of individuals at greatest risk, but amongst the much larger numbers of individuals who are at moderate (average) levels of absolute risk.[52] However, whilst it has been demonstrated that the effectiveness of the population strategy has probably been underestimated by the failure to take account of within–person variation,[49] little attempt has so far been made to examine the potential impact of current high–risk strategies and to compare these with the potential impact of population strategies. This leads to the question:

> (Q4) How does within–person variation in risk factors affect our understanding of the potential impact of different strategies for CHD prevention?

1.4.3 Impact on the understanding of social inequalities in CHD

Social inequalities in the incidence of CHD in the UK, with higher rates amongst lower social class groups, are well documented,[53;54] but while absolute CHD rates have fallen during the last 20 years, the fall has been concentrated among the highest social class groups so that the relative differences between those at the top and those at the bottom of the social scale have widened.[54] Consequently, recent CHD prevention policies in the United Kingdom have placed considerable emphasis on reducing these inequalities.[54;55] However, several important issues remain unresolved. First, little attention has been given to the observation that (like causally related risk factors), markers of socio-economic status are not always precisely estimated and may change over time. The effect of this on the size of estimated social inequalities has not been assessed. Furthermore, the relative contribution of adult and early life factors to social inequalities in coronary heart disease, as well as the likely value of identifying further factors influencing social inequalities, remains uncertain. Therefore:

(Q5) How does within–person variation in risk factors affect our understanding of social variations in CHD?

(Q6) How does this affect the estimated effectiveness of strategies aimed at reducing social inequalities in CHD?

1.5 The British Regional Heart Study: an opportunity to study the effects of within–person variation

The British Regional Heart Study (BRHS) is a prospective study of cardiovascular disease in 7,735 middle–aged men from 24 socially and geographically representative British towns. Study participants have had a wide range of risk exposures measured, and have been followed for fatal and non-fatal cardiovascular events since their enrollment into the study between 1978 and 1980, with very few losses to follow-up. The major aims of the BRHS are: (i) to determine the reasons for the geographic variation in cardiovascular disease in Britain; (ii) to investigate the individual determinants of cardiovascular disease; and (iii) to

study the occurrence, natural history and management of cardiovascular disease in Britain. This thesis aims to make a contribution to these aims by: (1) estimating the extent and effects of within–person variation in coronary risk factors on the estimated importance of the "established" coronary risk factors; (2) estimating the potential combined contribution of the strongest risk factors to population levels of CHD; (3) estimating the potential effectiveness of different strategies for the primary prevention of CHD; and (4) estimating the size of social class differences in CHD, the extent to which they may be explained by established CHD factors, and the implications this has for strategies aimed at reducing social inequalities in CHD.

Complete details of the BRHS methodology are presented in chapter 4. The baseline examination in 1978-80 consisted of a physical examination, an administered questionnaire (ascertaining social and lifestyle characteristics, history of previous disease, current CHD symptoms, and medication use), a non-fasting blood sample, and a resting electrocardiogram (ECG). Since the baseline examination, the men have been followed up for both fatal and non-fatal CHD events through National Health Service (NHS) central registers and biennial reviews of patients' General Practice medical records. A re-screening of surviving study participants after 20 years of follow-up (at which physical and biochemical measures were taken), together with similar examinations for men from two towns in 1996, as well as for an age-matched group of men over a one-week period in 2000, provides the necessary information on the effects of within–person variability in physical and biochemical risk factors. Furthermore, changes in lifestyle characteristics have been ascertained from postal questionnaires sent to surviving study participants after approximately 5, 13, 17 and 20 years of follow-up.

Two key features of the design of the BRHS are particular strengths with regard to addressing the objectives in this thesis. First, the BRHS is a representative population-based study and has accurate information on both fatal and non-fatal coronary heart disease events. Deaths from coronary heart disease were defined as all deaths with an International Classification of Diseases (ICD) code (9th revision) of 410-414 (including sudden death of presumed cardiac origin), while non-fatal myocardial infarction was defined according to established World Health Organisation (WHO) criteria. The primary outcome considered in this thesis is "major CHD", defined as coronary death or non-fatal myocardial infarction. A second advantage of the BRHS is that the study is purely obser-

vational - no attempt was made to influence clinical practice in the participating general practices.

1.6 Objectives and structure of the thesis

This thesis presents an epidemiological study of risk factors for coronary heart disease in British men, with an emphasis on the role that within–person variation in risk exposures has on disease relationships and prevention strategies. The aims of this thesis are listed below:

1. To estimate the extent of regression dilution bias in established and novel coronary risk factors over the 20-year study period, to assess whether the size of this bias varies with age, history of previous CHD, area of residence or social class, and to estimate the extent of within–person variation in "lifestyle" risk factors for CHD over the 20-year study period.

2. To examine the influence of regression dilution bias in continuous risk factors and within–person variation in lifestyle risk factors on the estimated age-adjusted (univariate) relationships between usual or average levels of these characteristics and major CHD risk over the 20-year study period.

3. To examine the potential combined contribution of the three strongest risk factors to major CHD risk (blood cholesterol, blood pressure and cigarette smoking) in individuals initially free from CHD, after correction for regression dilution bias.

4. To compare the potential effectiveness of high–risk and population approaches to the primary prevention of major CHD, after correction for regression dilution bias.

5. To estimate the size of social inequalities in CHD after adjustment for imprecision in the ascertainment of socio-economic status, to estimate the extent to which these differences may be attributed to differences in the established coronary risk factors, and to assess the potential effectiveness of strategies aimed at reducing social inequalities in CHD.

The content of each chapter is outlined below:

- **Chapter 1** gives a general introduction to coronary heart disease in the United Kingdom, outlines some of the difficulties of estimating risk relationships with CHD, and presents the thesis objectives.

- **Chapter 2** provides the epidemiological background to the research presented in this thesis, including a review of the evidence surrounding both the established as well as the new risk factors for CHD, a review of social and geographic inequalities in CHD in the United Kingdom, and a review of the different approaches to the primary prevention of CHD.

- **Chapter 3** provides the statistical background to the research presented in this thesis, including a review of the effects of within–person variation in risk factors on estimated disease relationships, and a summary of the particular statistical techniques and regression based approaches that are used in the thesis.

- **Chapter 4** describes the methods of the British Regional Heart Study, with particular attention given to the methods and definitions relevant to this thesis.

- **Chapter 5** is the first of five results chapters. The effects of regression dilution bias in established and novel coronary risk factors over a 20-year period are estimated, the extent to which they differ depending on previous CHD and social and geographic factors is assessed, and the consistency of the results with those of other similar studies is discussed. Furthermore, the extent of "within–person variation" in "lifestyle" characteristics over the 20–year follow–up period is assessed.

- **Chapter 6** uses these estimates of within–person variation to examine the effects that they have on univariate relationships with 20–year major CHD risk. The consistency of effects over the 20–year period is explored, and the combined (multivariate) effects of regression dilution in two factors measured with error and subject to variation over time are examined.

- **Chapter 7** examines the combined contribution of blood lipids, blood pressure and cigarette smoking to population levels of CHD among men initially free from CHD followed over the first ten years of the study (1978/80 - 1988/90).

- **Chapter 8** estimates the potential effectiveness of high–risk strategies to the primary prevention of major CHD compared with population strategies of prevention,

and is again based on men initially free from CHD followed over the first ten years.

- **Chapter 9** estimates the size of social inequalities in major CHD in men initially free from CHD followed over the first 20 years, and estimates the extent to which these differences are explained by established adult and early life factors.

- **Chapter 10** draws together the findings from this thesis and considers the implications for future epidemiological studies, future CHD research, and public health.

1.6.1 Choice of primary endpoint

The primary endpoint used in this thesis is the development of major CHD within 20 years (defined as non–fatal myocardial infarction or death from coronary heart disease; see Chapter 4 for details of case ascertainment). This endpoint for chosen in preference to a stricter definition of CHD (coronary death only) or an expanded endpoint (including a diagnosis of angina) in order to strike a balance between obtaining a hard (accurately defined) clinical endpoint on the one hand, while allowing for as many clinical events to be included in analyses as possible (for the sake of statistical power) on the other. In the results chapters of this thesis, the risk of major CHD over 20 years is primarily examined only in men initially free from symptoms or a diagnosis or CHD, in order to reduce the potential role that reverse causality bias may play. However, in chapter 6, analyses are based on all study participants (including those with pre–existing disease), in order to present a representative picture of the age–adjusted associations in the population of middle–aged British men between 1978/80 and 1998/2000. Furthermore, in chapters 7 and 8, 10–year follow–up data is used in analyses in preference to 20–year follow–up data. This was for a number of reasons: (1) the specific interest in considering preventable (i.e. early) major CHD in these chapters; (2) the desire to exclude the potential influence of the increased use of risk reducing drugs taken during the 1990's from results; (3) the ability to compare findings with current national primary prevention guidelines (which are based on 10–year CHD risk); and (4) the desire to adjust for regression dilution bias using real data observed over a four–year period, rather than estimating the "likely" effects of regression dilution bias at ten years (see chapter 5).

Each of the results in chapters 5 to 9 follow the same format: a summary of findings; a brief description of the background to the specific objectives in that chapter; the ob-

jectives and content of the chapter; a section detailing the methodology specific to that chapter; results; and discussion in which the validity and interpretation of the findings are considered in the context of the existing literature. Implications of the findings are not discussed in each individual chapter but are considered together in chapter 10. The thesis appendices include copies of the publications (to date) arising from this research (Appendix A), and the baseline questionnaire of the BRHS (Appendix B).

Figure 1.1: Age standardised CHD mortality rates in the United Kingdom between 1920 and 2000

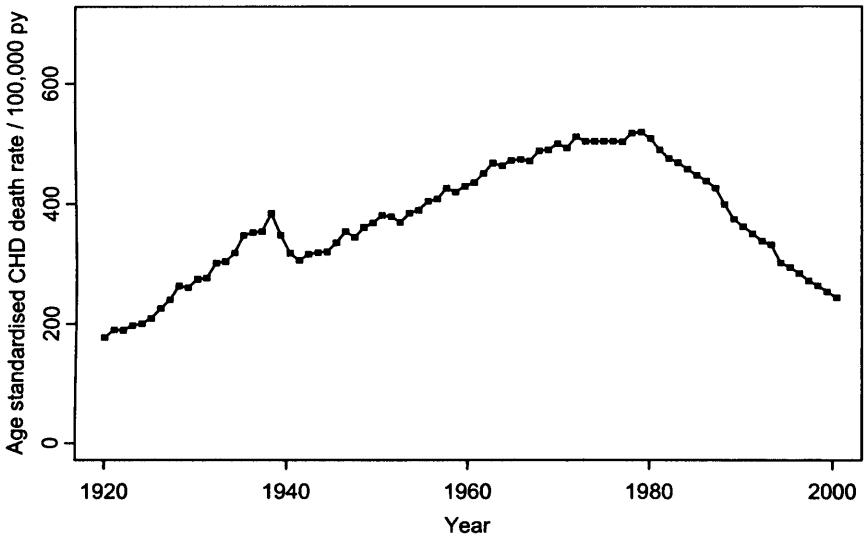
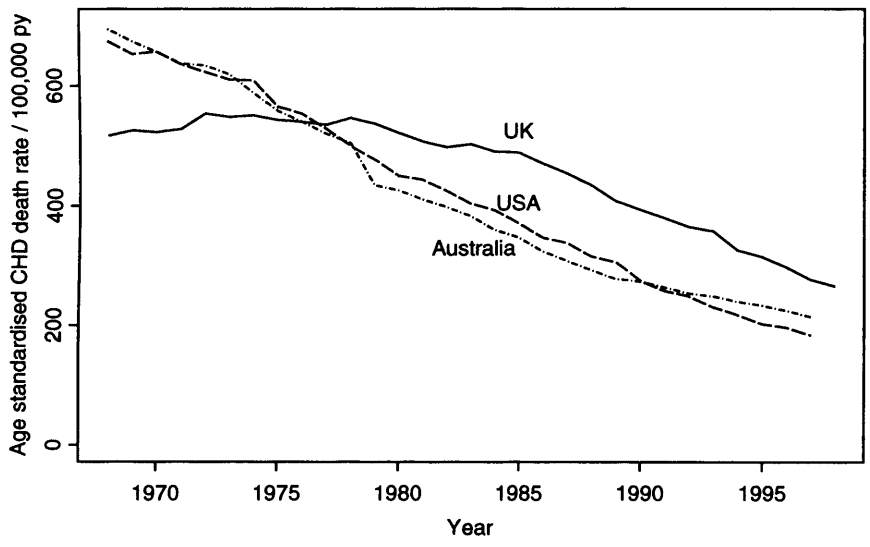


Figure 1.2: Age standardised CHD mortality rates in the UK, the USA and Australia between 1968 and 1998 in men aged 35–74 years (Source: World Health Organization statistics, 2001)



Chapter 2

Review of the epidemiology of CHD

2.1 Summary

Over the last fifty years, epidemiological studies have consistently demonstrated the importance of several risk factors for coronary heart disease, including adverse blood lipids, high blood pressure, cigarette smoking, physical inactivity, diabetes, and obesity. A combination of both observational and experimental evidence suggests that several of these are causal; randomised controlled trials clearly demonstrate the benefits on CHD risk of reducing blood cholesterol and blood pressure, while epidemiological studies show that smoking cessation, weight reduction and increased physical activity all reduce the subsequent risk of CHD. More recently however (as the processes underlying atherosclerosis have become better understood), there has been considerable interest in the potential role of new risk factors for CHD, including infections, and nutritional, inflammatory, haemostatic and genetic factors, as well as the possibility that CHD may be determined by "early life" factors. Interest in these new mechanisms has been fuelled by the widely held opinion that the "established" risk factors only explain about 50% of all CHD events that occur. Furthermore, geographic and social inequalities in CHD disease have been found to be only partially attributable to the established risk factors, further motivating the study of new disease mechanisms. The extent to which an individual's risk of CHD is determined by already known risk factors, and the overall importance of these factors to population levels of CHD, has important implications for the potential effectiveness of primary prevention

policies aimed at reducing CHD incidence in the United Kingdom.

2.2 Introduction

This chapter provides the epidemiological background to the major areas of study in this thesis. Section 2.3 introduces the various types of study that have helped determine the causes of CHD, in particular the prospective cohort study, and gives brief descriptions of the design and main findings of three of the earliest and most influential studies of coronary heart disease. In sections 2.4 and 2.5, the epidemiological and trial evidence for the role of the established and new coronary risk factors is presented, with emphasis given to the description of the mechanisms of action involved. The potential combined contribution of these factors to overall CHD risk is presented in section 2.6 and, in particular, a review of the claim that *"only half of CHD risk may be attributed to the established risk factors"* is provided. Section 2.7 reviews the different approaches to the prevention of CHD, both in individuals with and individuals without a history of diagnosed CHD, and describes the relative merits of these approaches. Section 2.8 describes the geographic and social inequalities in CHD risk that have been observed in many countries including the United Kingdom for many years, and the extent to which they are determined by known differences in the levels of the established risk factors.

2.3 Early Epidemiological Studies

As death rates from coronary heart disease in Western countries began to increase during the 20th century, it became clear that there were substantial variations in CHD mortality between different countries. This raised the possibility that coronary heart disease was a preventable disease, and not an inevitable consequence of ageing. Following the Second World War, several observational studies were initiated to investigate the possible reasons for these differences in CHD mortality between different countries. These tended to be either pathological studies (that investigated the determinants of pathologically defined coronary heart disease), or epidemiological (cohort) studies (that aimed to identify risk factors for the development of CHD in individuals). One of the earliest and most influential pathological studies was the International Atherosclerosis Project[56] which aimed to quantify the role of atherosclerosis in the development of coronary heart disease in differ-

ent populations. At about the same time, observational cohort studies were initiated to identify risk factors for CHD within a particular population. Part of the reason for this was because there was growing interest in the possibility of being able to identify individuals without symptomatic CHD but at high–risk of developing it. These cohort studies tended to be either cross–sectional studies (comparing risk factors with the prevalence of CHD at a single point in time) or prospective studies (comparing risk factors in individuals with the subsequent incidence of disease). In addition, some retrospective studies (looking at historical data) were also initiated. Of these types of study, the prospective cohort study has become the established design for identifying risk factors for CHD because, crucially, assessments of individuals are made before the CHD event occurs. Two of the earliest prospective cohort studies (the Seven Countries Study[29] and the Framingham Heart Study[28]) were particularly instrumental in identifying some of the causes of CHD, as well as helping to explain why CHD was common in some countries, but much less so in others. In particular, by collecting prospective data on individuals from several different countries, the design of the Seven Countries Study allowed the study of both the reasons for differences in CHD mortality between different populations as well as the study of the reasons for differences in CHD within particular populations (i.e. the causes of CHD at both the population and the individual level could be examined). In contrast, the Framingham Heart Study included participants from a single population, concentrating on assessing the determinants of CHD at the individual level, and in particular to develop methods for identifying individuals at greatest risk of developing CHD. The primary aims and findings from these studies, as well as those of the International Atherosclerosis Project, are now briefly described.

2.3.1 The International Atherosclerosis Project

The International Atherosclerosis Project[56] (carried out between 1960 and 1965) assessed atherosclerotic lesions and stenosis from 22,509 men and women who died aged 10-69 years from 14 countries in North, South and Central America, the Philippines, Jamaica, South Africa and Norway. The primary aims of the study were:

1. To assess the degree of atherosclerosis amongst individuals who died of causes unrelated to atherosclerosis.

2. To compare the level of atherosclerosis amongst communities with high coronary heart disease rates with those of communities with lower rates, and

3. To assess other circumstances (risk factors) that may aggravate the degree of atherosclerosis observed in individuals.

The study found that some degree of atherosclerosis is seen in all humans, but higher levels are consistently observed in communities where coronary heart disease rates are highest. At the population level, environmental conditions were found to be more important determinants of severe atherosclerosis than race or sex. In particular, the degree of atherosclerosis in a population was closely associated with both the proportion of total calories derived from fat and the average serum total cholesterol concentration in the population. Race did not appear to affect the level of atherosclerosis and substantial sex differences were only apparent in populations of high average atherosclerosis. The severity of atherosclerosis in a population was strongly related to the rates of clinical disease within that population, though other aggravating factors such as diabetes and hypertension further increased the likelihood of disease, particularly in populations with high levels of atherosclerosis. Within a population, individual levels of atherosclerosis varied considerably, suggesting that genetic factors influencing susceptibility to atherosclerosis at the tissue level (arteries) and the systemic level (lipid metabolism) may also be playing an important role.

The findings of the International Atherosclerosis Project have since been substantiated and extended in other pathological studies. In the Pathobiological Determinants of Atherosclerosis in Youth (PDAY) study, for instance, the percentage of intimal surface involved in atherosclerotic lesions in both the aorta and the right coronary artery was found to be positively associated with serum LDL and vLDL cholesterol, and negatively associated with serum HDL cholesterol, while serum thiocyanate (a marker for cigarette smoking) was strongly and independently associated with the prevalence of raised lesions.[57]

2.3.2 The Seven Countries Study

The Seven Countries Study[29] was the first study of its kind to examine the prospective relationships between lifestyle, diet, and the rates of cardiovascular disease in populations

with contrasting patterns of heart disease. The study (established by Ancel Keys in the early 1950's) consisted of 16 cohorts of men (initially aged 40–59) drawn from seven countries (the United States, Japan, Yugoslavia, Finland, Italy, the Netherlands and Greece) which were known to have widely differing levels of CHD mortality and which were suspected to have different dietary habits. Baseline data on more than 12,000 men were collected between 1957 and 1962, and individuals have been followed–up for cardiovascular events until the present day. Though the Seven Countries Study contains information on the relationship between specific risk factors and CHD in individuals (within each population), the primary focus of the study was to examine the reasons for international differences. In reports published after 5 and 10 years of follow–up, the study established that the key determinant of a population's risk of CHD was the mean serum total cholesterol concentration in the population, which was determined to a large degree by the percentage of total calories derived from saturated fats.[29] The low CHD rates observed in the Mediterranean countries of the Seven Countries Study gave rise to the suggestion that these countries were being "protected" from CHD by the composition of their national diets. Indeed, the Seven Countries Study is credited with discovering the "Mediterranean diet", as well as introducing the idea of mass causes and preventability of coronary heart disease.

2.3.3 The Framingham Heart Study

The Framingham Heart Study[28] was established in 1948 in Framingham, Massachusetts. The study was set up as a prospective observational study of approximately 5,000 men and women aged 30–59, and was initiated to investigate the strength of associations between lifestyle factors and physical and biological measurements with the risks of subsequent cardiovascular disease, in order to quantify the probability that any given member of a group of individuals at increased risk of CHD would suffer an episode of CHD within a certain time period. Physical and biochemical measurements were examined at a baseline assessment and re–examined every two years, with follow–up for cardiovascular mortality and morbidity continuing until the present day. Between 1968 and 1975, a second study comprising the offspring from the original Framingham study participants was established. These two Framingham studies (the Framingham Heart Study and the Framingham Offspring Study) have since become synonymous with the concept of "risk–scoring" through-

out the world.

Since the Framingham Heart Study was initiated, its design has been replicated in a vast number of subsequent prospective cohort studies, including further studies of American populations (e.g. the US Health Professionals study,[58] the Nurses' Health Study[59] and the Multiple Risk Factor Intervention Trial[60]), British populations (e.g. the British Regional Heart Study,[61] the Whitehall study,[62] the Renfrew/Paisley study[63] and the Scottish Heart and Health Study[64]), other European populations (e.g. the Paris[65] and Oslo[66] studies) and other non–European populations (e.g. the Shanghai study[67] and the Puerto Rico Heart Health Program[68]). Together, these studies have identified a vast number of risk factors associated with increased risk of CHD, as now described.

2.4 The established coronary risk factors

The early observational studies identified large differences in the risk of CHD between individuals, and the role of three important modifiable risk factors for CHD were quickly established: high blood total cholesterol, high blood pressure and cigarette smoking. These three risk factors were particularly relevant to the CHD epidemics developing in Western countries, not only because the relative risks associated with them were high, but because, crucially, they were widely distributed in these populations. Numerous subsequent studies have supported these early findings, and have also helped to establish other risk factors including physical inactivity, obesity, and diabetes, as important contributors to CHD risk. Together, these observational studies have contributed towards the wealth of evidence that now exists indicating that, at the population level, it is the lifestyle associated with "Western" cultures – a diet rich in saturated fats and calories, tobacco smoking and physical inactivity – that have important roles as causes of the mass occurrence of CHD in populations and as contributing factors to the risk of CHD in individuals within populations.[69] In the sections that follow, a description of each of these established risk factors is provided, with particular emphasis on the strength and consistency of the epidemiological evidence, the consistency between the epidemiological evidence and the evidence from experimental studies (trials), the mechanisms through which they influence CHD risk, and the factors that determine their levels.

2.4.1 Blood total cholesterol

Epidemiological and trial evidence

Cross–sectional studies, prospective cohort studies and clinical trials have all emphatically demonstrated that the risk of CHD increases with increasing levels of blood total cholesterol. This relationship has been shown to hold across populations at different CHD risks and populations with different average cholesterol levels, and has been shown to exist for both men and women, in individuals with or without pre–existing disease, and in individuals of all races. Furthermore, evidence from multiple prospective studies has demonstrated that the relationship is continuous with no "threshold" level below which a lower blood total cholesterol level does not confer a lower CHD risk[70–73] (see Figure 2.1). By pooling the findings from several prospective cohort studies, it has been estimated that, irrespective of initial cholesterol level, a 0.6 mmol/l lower serum cholesterol level is associated with a 54% lower risk of CHD at age 40 years, a 39% lower risk at age 50 and a 27% lower risk at age 60.[70]

Can the risk associated with a high blood cholesterol level be reversed? Early clinical trials of cholesterol reduction led to mixed results; randomised trials of the use of fibrates to lower cholesterol suggested that, despite reductions in CHD mortality, these drugs may increase the risk of total mortality.[74–76] In the World Health Organization trial of clofibrate, for instance, total mortality increased by 47% in the clofibrate group during treatment,[74] with the increase in mortality only levelling off a few years after clofibrate use was stopped. However, a parallel increase in non-cardiovascular deaths was not observed in dietary trials of cholesterol reduction (despite these trials showing similar reductions in CHD mortality[77]), suggesting that the observed increase in mortality from non–cardiovascular causes following treatment with fibrates may have been due to specific effects of the drug rather than to adverse effects following cholesterol reduction *per se*. The debate over whether cholesterol reduction increased the risk of death from non-cardiovascular causes was not definitively resolved, however, until the advent of a new class of lipid lowering drugs, 3–hydroxy–3–methylglutaryl coenzyme A (HMG–CoA) reductase inhibitors, or statins. Six large randomised placebo–controlled trials of these drugs (published between 1994 and 2002; see table 2.1) emphatically demonstrated that treatment with statins substantially reduced the subsequent risk of CHD and total mortality in both populations with[78–81]

and populations without[82;83] previous CHD. Furthermore, the largest of these studies (the Heart Protection Study)[78] further demonstrated these benefits to be similar in individuals without diagnosed coronary disease as in those who had cerebrovascular disease, peripheral artery disease, or diabetes, to be similar in men and women, in those aged under 70 years and those aged 70 or over, and, most notably, in those who presented with relatively low blood cholesterol levels (LDL cholesterol below 3.0 mmol/L or total cholesterol below 5.0 mmol/L) compared with those with higher levels. In a meta–analysis of the five trials that had been reported by 1999 (4S,[80] WOSCOPS,[82] CARE,[79] LIPID[81] and AFCAPS/TexCAPS[83]), it was concluded that statin drug treatment reduced major coronary event risk by 31% and all-cause mortality risk by 21%.[84] However, since the degree of risk reduction in these trials was predominantly determined by the level of cholesterol reduction achieved, greater reductions in risk are likely to be possible from statin doses that produce larger reductions in cholesterol.[73] This is supported by data from the first five statin trials (shown in Figure 2.2; reproduced from Ballantyne[85]), where the CHD event rates in the placebo and treatment arms of these studies have been plotted against the blood cholesterol levels achieved. In both primary– and secondary–prevention patients, it can be seen that the CHD event rate decreased with successively lower concentrations of low–density lipoprotein cholesterol achieved with either statin therapy or placebo (consistent with the observational relationship). This suggests that further reversal of CHD risk would have been possible had greater reductions in cholesterol been achieved.

How does blood cholesterol influence CHD risk?

The causal mechanism behind blood total cholesterol and CHD risk is complex, and only a brief review is provided here. Serum total cholesterol is made of various different components including low-density lipoprotein (LDL) cholesterol and high-density lipoprotein (HDL) cholesterol. The LDL component is the major atherogenic influence; increased LDL cholesterol can lead to the formation of lipid–rich plaques on the walls of the arteries increasing the risk of thrombosis, usually through plaque rupture or fissure (see section 1.2.2). In contrast, HDL cholesterol carries the LDL away from the arteries back to the liver, where it is excreted from the body. It is also thought that HDL removes excess cholesterol from plaques and hence slows their growth. HDL cholesterol is therefore often referred to as the "good" cholesterol, as high HDL reduces the risk of CHD. In

addition, the protein constituents of LDL and HDL cholesterol (known as the apolipoproteins) are increasingly becoming recognised as important markers of lipid metabolism and transportation, though these are not described further here.

Statins work primarily by targeting LDL cholesterol – they inhibit the HMG-CoA reductase enzyme in the liver (which controls the rate of cholesterol production in the body) slowing down the production of cholesterol and increasing the liver's ability to remove the LDL cholesterol component already in the blood. Though the extent to which the benefits of statins are due entirely to reductions in LDL cholesterol has been debated (it has been claimed that statins may have other beneficial "non–lipid" effects),[86;87] recent evidence from two of the statin trials indicated that virtually all of the treatment effects could be explained by the "on–study" lipid changes.[88;89]

Low blood cholesterol and all–cause mortality

In contrast to the continuous positive relationship between total cholesterol and the risk of CHD, the extent to which a lower serum total cholesterol level may lower the risk of all–cause mortality was (until the publication of the statin trials) hotly debated. Low blood cholesterol has been found to be associated with an increased risk of mortality from non–cardiovascular causes, particularly cancer, in numerous studies.[90–101] Most,[90–97] though not all,[98–101] have found this relationship to be explained, at least partially, by the preclinical effects of cancer on blood cholesterol levels. This is reflected by the observation that the cancer–cholesterol relationship tends to be markedly attenuated when events in the first few years of follow–up are excluded. In a systematic review carried out in 1994 of 10 large prospective studies, 2 international studies and 28 randomised trials, the authors concluded that there was:

> *"...no evidence that low or reduced serum total cholesterol increased mortality from any cause other than haemorrhagic stroke"*

but even then that:

> *"...the risks would be outweighed by the decreased risks from CHD."*[102]

These views have since been supported by the randomised controlled trials of statins; even in individuals with initially "average" cholesterol levels, statins significantly reduce all–cause mortality and have no adverse effects on cancer incidence.[78;103]

Determinants of blood cholesterol

The main determinants of serum total cholesterol concentrations in populations are dietary intake levels of saturated fat, polyunsaturated fat, and cholesterol.[104–107] Of these, saturated fatty acids are the most important; a diet high in saturated fats leads to increased serum total cholesterol. In contrast, polyunsaturated fatty acids lower serum cholesterol (monounsaturated fatty acids have no independent additional effects). Dietary cholesterol (which is found only in animal products) also increases levels of serum total cholesterol, though its effects are lower than those of the saturated fats. Cholesterol concentrations are also affected by reduced energy intakes resulting in weight loss[108] and possibly also by specific dietary supplements such as fibre,[109] garlic,[110] and fish oils.[111] Though diet is the most important contributing factor to average blood cholesterol levels in a population, it directly contributes only approximately 20% of the cholesterol present in the human body, the remaining 80% being produced by the liver. However, diets rich in saturated fats can lead to the liver increasing cholesterol production, and so the overall importance of diet on blood cholesterol is greater than these figures first suggest. In addition to diet, genetic factors are also likely to play a role in determining the variation in blood cholesterol levels observed between individuals.[112] A clear example of this is provided by familial hypercholesterolaemia, a condition affecting approximately 1 in 500 individuals (in Western countries), caused by a mutation in the LDL receptor gene and characterised by vastly increased cholesterol levels in the blood (see section 2.5.7).

2.4.2 Blood pressure

Epidemiological and trial evidence

As with blood total cholesterol, analysis of multiple prospective studies has emphatically demonstrated that the relationship between blood pressure and CHD (and stroke) is continuous with no thresholds below which risk does not continue to decrease (as least down to 75 mmHg for diastolic pressure and 115 mmHg for systolic pressure). In the Prospective Studies Collaboration, data on approximately one million individuals from 61 prospective studies revealed that a difference of approximately 20 mmHg in systolic pressure or 10 mmHg in diastolic pressure corresponds to a more than two–fold difference in the risk of fatal CHD during middle-age[113] (see Figure 2.3), with this relative risk difference atten-

uating with increasing age. Furthermore, analysis of approximately 60,000 non-vascular deaths from the same study has shown that there are no appreciable risks associated with low blood pressure.[113]

Lowering blood pressure reduces CHD risk, irrespective of how this is achieved.[114–120] This can be seen in Figure 2.4, for example, where the effects of different blood pressure lowering drugs on cardiovascular mortality risk from 27 randomised controlled trials are shown.[120] The most common classes of blood pressure lowering drugs are diuretics, β–blockers, calcium channel blockers and angiotensin converting enzyme (ACE) inhibitors. Interestingly, these drugs lower blood pressure in different ways: diuretics assist the kidney to excrete fluids and salt and can also help to widen blood vessels, β–blockers block adrenaline thus slowing the heartbeat, calcium channel blockers block the calcium needed for muscle contraction and hence reduce arterial or heart muscle tension, while ACE inhibitors interrupt the formation of angiotensin II which makes blood vessels contract. To assess the relative effectiveness of different blood pressure lowering drugs, the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT) was initiated. This trial included over 33,000 high risk hypertensive men and women aged 55 or over who were randomised to receive either chlorthalidone (a diuretic), amlodipine (a calcium channel blocker) or lisinopril (an ACE inhibitor). The primary results of the ALLHAT study (reported in 2002) showed that over 5 years of follow–up, there were no differences in CHD or all–cause mortality risk between the amlodipine and chlorthalidone treatment groups or between the lisinopril and chlorthalidone groups.[121] These findings were subsequently confirmed in an overview of 29 randomised controlled trials (including ALLHAT) published in 2003.[116] The authors from the Blood Pressure Lowering Trialists' Collaboration reported no significant differences in total major cardiovascular events between treatment regimes based on ACE inhibitors, calcium channel blockers, diuretics or β–blockers, and found that for every outcome other than heart failure, the difference between randomised groups in achieved blood pressure reduction was directly related to the observed difference in disease risk. By combining several blood pressure lowering drugs together, greater reductions in blood pressure (and hence greater reductions in CHD risk) are likely to be achievable than with any single drug in isolation.[115;122] A recent analysis of 354 randomised trials of blood pressure lowering agents examined this hypothesis and found that the blood pressure lowering effects of different categories of drugs were addi-

tive. The authors estimated that a combination of three drugs at half standard dose (for example, a diuretic, a beta blocker and an ACE inhibitor) would lower blood pressure by 20 mmHg systolic and 11mm Hg diastolic, approximately three times that of any single drug at half dose, or twice that of any single drug at standard dose, and would also result in fewer side effects.[122]

How does blood pressure influence CHD risk?

Given the vast amount of evidence relating blood pressure to CHD risk, as well as the number of clinical trials that have demonstrated CHD risk to be reversible through blood pressure lowering drugs, relatively little is actually known about the mechanism through which blood pressure influences CHD risk. However, it is thought that the effects may, at least partially, be mediated through inflammatory processes (see section 2.5.3), possibly through the formation of free radicals, as well as through haemodynamic effects.

Determinants of blood pressure

One of the most important determinants of the average blood pressure level within a society, as well as differences in blood pressure between societies, is dietary salt intake.[123–126] At age 60–69, it has been estimated that a difference in sodium intake of 100 mmol per day (approximately 6 grams of salt a day) is associated with an average difference in systolic blood pressure of around 10 mmHg,[126] and an average diastolic blood pressure difference of around 5 mmHg. These associations are similar in countries with high average blood pressure (usually developed countries) and countries with low average blood pressure, and are consistent with the observed effects on blood pressure reported in salt reduction trials.[127] However, dietary salt is by no means the sole determinant of an individual's blood pressure level. Factors such as physical inactivity, increased body mass index,[128] a diet low in fruit and vegetables and high in saturated fat,[129] low potassium intake[130] and high alcohol intake[131] are all associated with increased blood pressure independently of dietary salt intake. Low birth weight has also been found to be associated with increased adult blood pressure levels,[132–135] though the strength of this relationship is more controversial.[136] Furthermore, it is possible that since black individuals and those with a family history of high blood pressure are more likely to have high blood pressure,[137] genetic factors may play some role in determining blood pressure levels.

2.4.3 Cigarette smoking

Once the link between cigarette smoking and lung cancer was established in 1950[138;139] (a finding later confirmed by both the Seven Countries Study and the Framingham Heart Study), observational cohort studies in Britain and the USA were initiated to explore the prospective relationships not only with lung cancer, but also with deaths from other causes, including CHD.[140;141] The British study consisted of 40,000 doctors (the British Doctors Study) and initial results published in 1954 demonstrated for the first time that cigarette smokers were at a substantially greater risk of dying from CHD than non–smokers.[142] Subsequent reports after ten,[143] twenty,[144] forty[145] and now fifty[146] years have reiterated these findings and have shown that though the relative risk of fatal CHD for smokers relative to non–smokers is smaller than that observed for the various forms of cancer more closely related to smoking (e.g. lung cancer), the greater number of deaths attributable to vascular causes meant that the absolute excess mortality from vascular diseases in cigarette smokers was more than double that attributed to cancers. Indeed, cigarette smoking is currently estimated to kill over two million people a year in developed countries (half during middle age), which accounts for approximately one sixth of the adult deaths each year in these populations.[147;148] Of these, over half are due to cardiovascular disease. More recent studies have further established that this adverse effect of smoking is related both to the amount of tobacco smoked and the duration of smoking.[149;150] A study of 14,000 survivors of myocardial infarction (cases) and 32,000 of their relatives (controls) carried out in 1995 further identified the importance of tar yields in the development of non–fatal myocardial infarction.[151] In addition, passive smoking has now also been shown to increase the risk of CHD and other smoking related diseases.[152–154] Non–smokers who live with smokers may have up to a 30% increased risk of CHD over non–smokers who are not exposed to environmental tobacco smoke,[153;154] though the true risks associated with all passive smoking (not just smoking in the home) may be even greater.[155;156]

Mechanisms and effects of smoking cessation

The exact mechanisms through which tobacco smoking increases the risk of atherosclerotic disease are not yet fully understood, though it is thought that they are multiple. In particular, cigarette smoking is thought to affect both atherosclerotic and thrombotic (clotting) processes, is thought to damage the endothelium through free radical formation,

to affect platelet formation and aggregation, and even to have adverse effects on blood lipids.[157–162] However, regardless of the true nature of the mechanism, it is well established that stopping smoking, even during middle age, substantially reverses the risks of CHD and death.[144;145;163–167] Data published in 1992 from the British Regional Heart Study, for instance, indicated that middle–aged men who had quit smoking for more than five years did not differ appreciably in their risk of CHD over the following 10 years from men who had never smoked cigarettes.[168]

2.4.4 Physical inactivity

Physical inactivity was first regarded as a potential contributing factor to CHD in the early years following the Second World War when Professor Jeremy Morris and his colleagues observed that bus conductors (whose jobs required them to be physically active) had approximately half the CHD risk of that of bus drivers (whose jobs were more sedentary).[169] However, at the time, there was much scepticism about the role of physical activity in CHD risk, the conventional thinking was that CHD resulted from hypertension, hypercholesterolaemia and obesity. Indeed it was not really until the publication of results from two large studies of middle–aged British civil servants (initiated in 1968 and 1976),[170;171] as well as a study of approximately 17,000 Harvard alumni (followed between 1962 and 1978),[172] that physical inactivity became widely established as an independent risk factor for CHD.

Since these early studies, numerous subsequent studies have demonstrated the protective effects of physical activity on the risk of CHD[173–177] leading to it becoming universally recognised as a protective factor for CHD. Though many of these studies were based on middle–aged men, the benefits of physical activity on the risk of CHD have also been substantially demonstrated for women.[178–182] Furthermore, physical activity has been shown to reduce CHD risk in specific groups of high–risk individuals, including the elderly,[179;183–186] men with established CHD[187–189] and individuals with diabetes.[190] However, despite this wealth of evidence, there have been conflicting findings regarding the type, frequency and intensity of physical activity that is needed to achieve benefit. The influential study of middle–aged British office workers by Morris and colleagues found that in order to achieve reductions in CHD risk, leisure time physical activity needed to consist of vigorous aerobic activity at least twice a week, and that recreational work such

as walking, gardening and "Do–It–Yourself" (DIY) did not provide protection.[171] These findings were later supported by studies of Finnish[191] and German[192] men, which helped develop the view that in order to be beneficial, leisure time physical activity needed to be frequent and vigorous in nature. This viewpoint has been questioned in recent times, with many studies demonstrating that even light to moderate levels of physical activity significantly reduced the risks of developing CHD.[176;186;193–195] In addition, the uptake of physical activity among sedentary individuals, even in later life, has been shown to confer significantly lower risks of CHD and all cause mortality.[196]

Mechanisms

The strength and consistency of the relationship between physical activity and CHD strongly suggests a causal relationship, though mechanisms are still unclear. Physical activity is associated with several other coronary risk factors, including blood pressure[197] and blood lipids.[198;199] Therefore the relationship between physical activity and CHD could be mediated at least partially through these risk factors. However, most studies have found associations between physical activity and CHD to be independent of these risk factors. Alternatively, since physical activity reduces insulin resistance (see section 2.4.6),[200;201] insulin and other components of the "insulin resistance syndrome" may be influencing the physical activity–CHD relationship. However, to date, few prospective studies have examined this possibility, and one that has found no evidence in favour of this hypothesis.[202] A further suggestion is that physical activity may reduce CHD risk through its effects in the acute phase of CHD, possibly through clotting processes, platelet aggregation or increased fibrinolytic activity (see section 2.5),[203–205] though this has not yet been conclusively established.

2.4.5 Diabetes

Insulin (a hormone released by the pancreas in response to increased levels of sugar in the blood) is important in order for glucose (blood sugar) to be absorbed into the body's cells. Diabetes mellitus is a disease defined by increased levels of glucose in the blood caused by a relative or absolute deficiency of insulin. There are two forms of diabetes: type I, or insulin dependent diabetes, and type II, or non-insulin dependent diabetes. Type I diabetes is a disease of the pancreas gland whereby the body is deficient in insulin, leading to increased

glucose levels in the bloodstream. Affected individuals need insulin injections in order to keep blood glucose levels as close to normal as possible. Type II diabetes generally occurs because of a metabolic failure at the cellular level – cells become "resistant" to the effects of insulin, and the concentration of glucose in the blood, unable to enter the cell, begins to rise. The body's natural response to insulin resistance is to secrete more insulin from the pancreas until the glucose is taken up by the cells. However, if the pancreas cannot sustain this state of hyperinsulinaemia (high insulin levels in the blood), blood glucose levels will increase leading to type II diabetes. Typically, individuals with type II diabetes do not require insulin injections as the diabetes can usually be controlled through dietary or pharmacological means.

Both type I and type II diabetes are associated with a markedly increased risk of CHD and other vascular diseases.[206;207] Prospective epidemiological studies on large cohorts of diabetic patients have shown that the degree of hyperglycaemia (high blood glucose) is positively associated with the subsequent risk of CHD.[208–210] However, in a large multicentre trial of Type II diabetics (the United Kingdom Prospective Diabetes Study[211]), there were no apparent effects of glucose control on the risk of coronary or all-cause mortality over ten years of follow–up.[212] Conventional coronary risk factors (e.g. high cholesterol, high blood pressure, cigarette smoking) appear to have the same relative impact in diabetics as non-diabetics, and so the absolute CHD risks of diabetics are (at any given risk factor levels) considerably greater than those for non-diabetics.[213;214] Several randomised controlled trials and meta–analyses have confirmed that the benefits of reducing blood cholesterol[215–217] and blood pressure[218–220] extend to patients with diabetes. Therefore, the absolute benefits from risk factor modification, including smoking cessation, in diabetic patients are particularly great. Given the increased CHD risks experienced by diabetic patients, it is natural to wonder whether lesser degrees of disturbance to the glucose and insulin metabolism also lead to increased CHD risks. These questions are now considered further.

2.4.6 Obesity, insulin resistance and the metabolic syndrome

Obesity and CHD

In the United States (and, increasingly, in the United Kingdom) obesity is becoming a national epidemic.[221] Although it is well established that obesity is an important determi-

nant of raised blood pressure[222] and type II diabetes,[223] the evidence for obesity as a risk factor for CHD, once blood pressure and blood lipids are taken into account, has been less well established. However, results from large prospective studies in both men and women have demonstrated that the risk of CHD does increase progressively as body mass index increases, that this increase is independent of blood pressure and blood cholesterol and that it begins at even moderate levels of weight gain and overweight.[224–230] In middle–aged American men and women, CHD risk has been shown to be as much as three to four times higher in those who are obese (a BMI of at least 30 kg/m^2) compared with those who are lean (a BMI of $<23 kg/m^2$).[224;228] A recent meta–analysis of the effects of body mass index on CHD risk in Chinese adults estimated a 7% increase in risk per 1 kg/m^2 increase in BMI, indicating a slightly lower (twofold) difference in CHD risk between those with a BMI of 20 and those with a BMI of 30.[229] Though most attention has been based on the role of general adiposity (as measured by body mass index), numerous studies have now indicated that central adiposity (defined as an increased intra–abdominal fat mass and measured by the "waist–hip" ratio) may be more strongly related to the risk of CHD than BMI, particularly in women.[228;231–233] In a study of approximately 33,000 women aged 55 to 69, nearly a threefold gradient in CHD mortality was observed between those in the lowest and highest tertiles of waist–hip ratio, and this was independent of other risk markers including body mass index.[233] In contrast BMI, showed no independent relation to CHD mortality in this population.

Insulin resistance, the metabolic syndrome and CHD

As well as increasing CHD risk, obesity, and in particular central obesity, is an important factor in the development of "insulin resistance" (the slow uptake of glucose into the tissues; see section 2.4.5). Although it is difficult to differentiate the relative effects of insulin resistance and hyperinsulinaemia as distinct entities, it is clear from a large body of basic, animal and human studies that insulin resistance is associated with significantly increased CHD risk, due in part to its associations with high blood pressure, high triglycerides and low HDL cholesterol.[234–238] In fact, the close nature of the inter-relationships between central obesity, insulin resistance, high blood pressure and adverse blood lipids has led to the four conditions becoming known as a syndrome referred to as the metabolic syndrome, syndrome X or the insulin resistance syndrome.[239] The interest in the metabolic syndrome

lies not just in the observation that individuals with the syndrome are at higher risk of diabetes[240] and CHD[241] than individuals without the syndrome (as would be expected because, with the possible exception of insulin resistance, the traits of the syndrome are well established risk factors for these diseases), but also in the hypothesis that the metabolic syndrome may increase risk for adverse outcomes to a greater degree than predicted by the individual components. This hypothesis is supported by a recent analysis that indicated that the syndrome traits interact to increase atherosclerosis of the carotid artery by a greater degree than would be expected solely from their additive effects.[242] Furthermore, secondary analyses of two cholesterol lowering trials found that among patients allocated to the placebo arm, the excess risks of major coronary events experienced by those with the metabolic syndrome were over and above that which could be explained by the traditional risk factors.[243]

2.4.7 Gender and hormone replacement therapy

In addition to being the single most common cause of death in men, coronary heart disease is also the most common cause of death in women, accounting for 13% of all female deaths (approximately seven times the number that are caused by breast cancer).[1] However, at any given age, women have a substantially lower risk of CHD compared with men.[244;245] For years, it was thought that the excess risk experienced by men was explained by unhealthy behaviours that were more socially acceptable for men than women (such as cigarette smoking, heavy alcohol use, and a poorer diet) rather than being caused by inherent sex differences in physiology.[246] However, studies adjusting for such behaviours generally found that while these factors contributed to the observed difference in CHD risk by gender, they did not fully explain the increased risk of CHD in men.[247] Subsequently, considerable attention was drawn to the potential role of oestrogen as a protective factor in premenopausal women, a hypothesis prompted in part by the observation that the relative differences in CHD incidence and mortality between men and women decreases with increasing age. This, in turn, led to the suggestion that hormone (oestrogen) replacement therapy (HRT) after the menopause could have a potential cardioprotective role. Observational studies appeared at first to support this view; in a review of epidemiological studies of the effect of postmenopausal oestrogen on coronary heart disease risk in 1991, Stampfer and Colditz[248] estimated that the relative reduction in CHD risk associated

with ever taking HRT compared with never taking HRT was 50% (95% CI 44% to 57%), concluding that, *"...the bulk of evidence strongly supports a protective effect of oestrogens that is unlikely to be explained by confounding factors..."*. In 1998 however, the first large, randomised, placebo–controlled trial of the effect of combined oestrogen/progestin hormone replacement therapy on coronary events – the Heart Progestin/Estrogen Replacement Study (HERS) – reported no beneficial effect of combined therapy on CHD events in women with established coronary disease.[249] In 2002, the Women's Health Initiative (WHI) trial of oestrogen plus progestin in 16,608 healthy postmenopausal women found that the relative risk of CHD was actually 29% higher (95% CI 2% to 63%) among those on active therapy compared with those on placebo (the combined therapy arm was actually stopped early by the data and safety monitoring board because of an increased risk of breast cancer).[250] In 2004, the oestrogen only arm of the WHI trial was also terminated early after showing that oestrogen alone increased the risk of stroke and had no effect on the risk of CHD.[251]

Why therefore, is there such a discrepancy between the findings from the randomised controlled trials and the observational studies? In contrast to the 1991 claim of Stampfer and Colditz,[248] it appears that residual confounding may provide the answer after all. In particular, HRT use is strongly related to socio–economic status throughout the life course,[252] which in turn is strongly related to CHD risk (see section 2.8.1). Regardless of the full reasons behind these differences however, the HRT–CHD relationship provides a valuable reminder as to why evidence from observational studies (even very good ones) should, wherever possible, always be interpreted in the context of supporting evidence from randomised controlled trials, and should not necessarily be accepted as fact.[253;254]

2.4.8 Alcohol

The relationship between alcohol and CHD risk (and total mortality) has consistently been shown to be U- (or J-) shaped, with individuals who drink light to moderate amounts of alcohol generally being around 25–30% less likely to experience CHD than individuals who do not drink, and individuals who drink excessively to be at increased risk of CHD (as well as other diseases).[255–257] This relationship has been consistently demonstrated for both men and women in a large number of prospective studies across many different populations.[258–264] But the issue of whether alcohol is causally protective at low doses has

been difficult to establish from prospective studies alone, due in part to the difficulties of comparing non–drinkers (a group consisting of a mixture of ex–drinkers and lifelong teetotallers) with regular occasional or light drinkers. Previously, it has been suggested that the protective effects of alcohol were due to a combination of adverse social and lifestyle characteristics among non–drinkers with particularly favourable characteristics of regular light/moderate drinkers and not due to alcohol itself.[265] However, in recent times, many studies have shown the benefits of alcohol to be independent of pre–existing disease, other coronary risk factors, and to be present even after exclusion of events early on in the follow–up period.[266;267] It is therefore now widely accepted that a moderate level of alcohol intake does protect against CHD risk. Several questions remain however.

1. **How does alcohol influence CHD risk?**

   Current opinion on this matter suggests the effects are mediated primarily through effects on lipids and fibrinolytic activity.[268] In particular, it has been estimated that between 40 and 60% of the protective effect of alcohol may be attributed to increases in HDL cholesterol[269–272] (though the true figure may be somewhat larger due to measurement errors in recording alcohol consumption and biological variability in HDL cholesterol). In addition, alcohol appears to improve fibrinolytic activity (see section 2.5.4) by reducing platelet aggregability[273;274] and reducing fibrinogen and factor VII levels.[275;276]

2. **Does the type of alcohol matter?**

   Whether or not different risks and benefits are associated with different types of alcoholic beverages is an important issue. Of particular interest is the possibility that substances in wine may have beneficial effects in addition to those of ethanol. Indeed, it has often been suggested that despite high smoking rates and high fat diets, the French experience low rates of cardiovascular disease because of their high levels of wine intake.[274] Numerous studies have therefore aimed to assess the specific effects of wine, beer and spirits on CHD risk. In populations where all three types of drink are commonly consumed, these studies have indeed often found wine drinkers to be at lower risk of CHD than beer or spirit drinkers.[277–279] A recent meta–analysis of 26 such studies estimated that wine reduced the risk of CHD by 32% whereas beer reduced this risk by 22%.[280] However, as yet, it remains unclear

as to which substances in wine may be having effects in addition to those of ethanol, though many have been suggested.[281–283]

3. **Does the pattern of drinking matter?**
   Drinking pattern has been shown to be predictive of CHD risk independently of average volume intake;[266;284] several studies have found that episodic consumption of large amounts of alcohol (binge drinking), rather than reducing CHD risk, leads to increased risks of CHD and all–cause mortality.[285–288] Whether or not this is due to differential effects on HDL cholesterol is unclear,[285;289] though it may be due, in part at least, to the different effects that drinking pattern can have on blood pressure.[290–292] The Intersalt study, for instance, found that a highly variable pattern of alcohol consumption predicted a high mean blood pressure among heavy drinkers, regardless of the amount of alcohol consumed the day before measurement.[131] Furthermore, it is thought that platelet aggregation may be increased among heavy drinkers during periods of withdrawal.[293]

Does the evidence about the benefits of regular light to moderate drinking suggest that non–drinkers should be advised to begin drinking? In the absence of clinical trial evidence, it would be difficult to recommend drinking alcohol as a preventive measure, particularly given the myriad of other health and social problems associated with alcohol, the risks associated with heavy drinking and the observation that many non–drinkers abstain for a particular reason (e.g. religion, previous health conditions, family history of alcoholism). However, without encouraging alcohol use, it seems reasonable to state that moderate alcohol intake can form part of a healthy lifestyle.

2.5 The novel risk factors

As the mechanisms behind atherosclerosis and CHD have become better understood, a variety of new hypotheses have been generated that implicate the potential importance of a number of novel risk factors for CHD. Interest in these factors has been driven by the widely held belief that the established risk factors described so far are only partially responsible for determining CHD risk.[30–40] In this section, a broad review of these new areas of research is provided; the background for each research hypothesis is briefly described and the evidence supporting the hypothesis reviewed.

2.5.1 Homocysteine

Homocysteine is a sulphur–containing amino acid present in all cells which, in the normal course of events, is broken down into other non–damaging amino acids by three B vitamins: folate, B–6 and B–12. Plasma homocysteine levels reflect both environmental and genetic factors,[294;295] and are inversely related to the dietary intake of folate and, to a lesser extent, vitamin B–12.[294] Individuals with the rare autosomal recessive condition homocystinuria have extremely high blood levels of homocysteine (>100 μmol/l), and a high incidence of vascular disease (approximately half of homozygotes are affected by the age of 30).[296] Studies of such individuals who died have prompted the hypothesis that even moderately elevated blood concentrations of homocysteine may be relevant to cardiovascular disease.[297]

Observational and trial evidence

The first studies to address this hypothesis were case–control studies of patients with CHD. These studies did report higher blood homocysteine levels in cases than in age matched controls who were free of disease, though the studies were small.[298;299] Since these early reports however, the amount of evidence from prospective studies has increased substantially. In 2002, a meta–analysis of 30 observational studies involving over 5,000 CHD events found that a 25% lower homocysteine level was associated with an 11% lower CHD risk (95% confidence interval 4 to 17%), and that this reduction was independent of other established coronary risk factors.[300] Results from genetic studies published the same year supported these findings, strengthening the view that the relationship is causal.[301] This is particularly relevant as homocysteine levels can easily be reduced by taking folic acid and other B–vitamins. A meta–analysis of 12 randomised trials of homocysteine reduction found that dietary supplementation with folic acid reduced blood homocysteine by approximately one quarter (independently of dose), and that further supplementation with vitamin B-12 produced an additional 7% reduction.[302] In the United States, folic acid has been added to flour since 1997, resulting in a significant reduction in average serum homocysteine.[303] In the late 1990's, several large-scale randomised trials in people with prior CHD, prior stroke or renal disease were initiated to test the hypothesis that homocysteine–lowering with folic acid could reduce the risk of recurrent cardiovascular disease.[304] Two of these studies were ended prematurely; the Second Cambridge Heart Antioxidant Study

(CHAOS-2) of 1,882 people with prior CHD was terminated early because of a perceived lack of power to address the hypothesis (although it was reported that lowering homocysteine by 13% had no significant effect on any vascular outcome),[305] while the Vitamin Intervention for Stroke Prevention (VISP) study of 3,600 people with prior stroke was terminated when the interim analyses showed little possibility of demonstrating any effect of treatment on vascular events.[306] The remaining trials should be completed from mid 2005 onwards.

Mechanisms

The mechanism through which homocysteine can increase the risk of CHD is thought to be due to the damage it can cause to the endothelium and its tendency to promote the buildup of LDL cholesterol, leading to blockage of the arteries. Genetic factors also modulate total plasma homocysteine levels, in particular the thermolabile variant of methylenetetrahydrofolate reductase (MTHFR),[307] the presence of which may be a risk factor in folate–depleted individuals (see section 2.5.7).[308] However, until the largest homocysteine lowering trials are reported, it is uncertain whether supplementation of foods with folic acid should be recommended for CHD prevention, though careful attention to conventional risk factors in individuals with a raised plasma total homocysteine is warranted regardless.

2.5.2 Antioxidants

When LDL crosses the endothelial membrane to enter the artery wall, it becomes subject to a number of modifications, including an oxidation process which results in the production of free radicals. Free radicals are atoms or molecules with an odd number of electrons (making them unstable, short lived and highly reactive). As they combine with other molecules with unpaired electrons, new free radicals are produced, leading to a chain reaction that can damage important cellular components, including the cell membrane. The body's natural response to this damage can lead to the creation of macrophages that take up the LDL before being converted to foam cells (see section 2.5.3 for further details). Indeed, oxidised LDL is thought to play a key role in the inflammatory processes that underpin the development of atherosclerosis. For these reasons, there has been much interest in the potential role that antioxidants (the body's natural defence against free radicals) may have in CHD prevention.[309]

Evidence from observational studies and clinical trials

Many epidemiological studies have shown that individuals whose diets are rich in the principal antioxidants (vitamin C, vitamin E and β–carotene) experience lower risks of cardiovascular disease.[310–318] A systematic review of all such studies published up to March 2001 (20 studies) estimated that a high intake of β–carotene or ascorbic acid (compared with a low intake) reduced the risk of CVD by approximately 11%, while a high intake of α–Tocopherol reduced this risk by 26%.[319] Furthermore, a recent pooled meta–analysis of ten American and European prospective studies found that after adjustment for demographic factors, body mass index and lifestyle factors, a 10 gram/day increase in total dietary fibre was associated with a 14% reduction in all coronary events and a 27% reduction in coronary death.[320] However, these findings from epidemiological studies have generally not been supported by evidence from randomised controlled trials in either individuals with,[321–325] or individuals without[326–331] previous evidence of CVD. For studies published before March 2001, there were no significant differences in the risks of cardiovascular disease between individuals randomised to receive β-carotene, α–Tocopherol or ascorbic acid relative to individuals randomised to placebo.[319] In 2002, the very large MRC–BHF Heart Protection Study confirmed these findings, by reporting no effect of vitamin supplementation on the 5–year risk of vascular (or cancer) mortality (despite the vitamin supplements leading to substantial increases in blood vitamin levels).[322] In 2003, a summary of all cohort and trial evidence on vitamin supplementation and cardiovascular disease (reported to the US Preventive Service Task Force) stated that:

> *"Despite promising evidence from cohort studies, randomised controlled trials have failed to demonstrate a consistent or significant effect of any single vitamin or combination of vitamins on incidence of or death from cardiovascular disease."*[332]

The discrepancy between the observational studies and the randomised controlled trials may be due to residual confounding in the observational studies, i.e. unknown or unmeasured factors which influence CHD risk but are also related to the level of fruit and vegetable intake in individuals, though another explanation is that the beneficial effects of antioxidants need to accrue over a period of many years.[333]

2.5.3 Inflammation and Infection

The inflammatory response by the body to injuries to the endothelium has long been suggested as an important factor in the process of atherogenesis.[334] There are many ways in which the endothelium can become damaged. However, whatever the cause of endothelial injury, atherosclerosis appears to be a characteristic response of particular arteries.

Inflammation, infection and atherosclerosis: pathological evidence

When LDL cholesterol particles become trapped in an artery they become subject to a variety of modifications (of which oxidation is the most well known; see section 2.5.2). These modified LDL particles are proinflammatory, attracting monocytes into the subendothelial space and promoting the differentiation of monocytes into macrophages (large white cells that protect against infections and toxins). This is a key step in the inflammatory process because macrophages take up LDL through scavenger receptors, accumulate the lipid, and are converted to the foam cells that form the basis of atherosclerosis. Though cholesterol is the inflammatory agent most regularly present in the walls of the arteries, other causes of endothelial dysfunction have been proposed. In particular, it has been suggested that certain infections may aggravate the inflammatory processes already present in the arterial walls (though perhaps without initiating the processes themselves). *Chlamydia pneumoniae* and *helicobacter pylori* have been identified as the likeliest infectious microorganisms that may contribute to the development of atherosclerosis and CHD. However, though there is no direct evidence to suggest that these can cause the lesions of atherosclerosis,[335] this possibility has become an area of widespread interest.

Epidemiological and trial evidence

Chlamydia pneumoniae infection was first proposed as an avoidable cause of coronary heart disease in a small retrospective study in 1988.[336] Subsequent studies suggested a twofold or larger odds ratio for coronary heart disease in people with markers of chronic *C pneumoniae* infection, though these studies tended to be small, retrospective, subject to confounding and liable to biases.[337–339] In comparison, in 2000 a meta–analysis of 14 prospective studies consisting of over 3,000 cases found that after adjustment for smoking and social class, only a small and insignificant association between *C pneumoniae* infection and CHD risk remained.[340] In addition, the contribution of *helicobacter pylori* to CHD

risk has been shown in prospective studies to be only weak and possibly due to social confounding.[341;342] C-reactive protein however, a marker of the inflammatory response to tissue damage, has consistently been shown to be prospectively related to CHD risk,[343–345] independently of traditional risk factors. In a meta–analysis of 7 long-term prospective studies of CHD published before 1998, a difference in C reactive protein of 1.4 mg/L (from 1.0 mg/L to 2.4 mg/L) corresponded to a 70% increase in the risk of CHD.[343] Additionally, in 2003, an 8–year follow–up of 14,000 initially healthy women found that differences in CHD risk by CRP level were similar in magnitude to those based on having at least 3 components of the metabolic syndrome and, furthermore, that CRP added prognostic information on subsequent CHD risk at all levels of severity of the metabolic syndrome.[344] These findings may have overestimated the true importance of C–reactive protein however because of the preferential publication of positive results in earlier studies. In a recent updated meta–analysis of 22 prospective studies, those in the top tertile of the C–reactive protein distribution had a 58% higher risk of CHD than those in the lowest tertile, which was reduced to 49% when only the four largest studies (4107 cases) were used.[346] However, randomised controlled trials have failed to demonstrate any effectiveness of antibiotic treatment on coronary risk. In the Weekly Intervention With Zithromax for Atherosclerosis and Related Disorders (WIZARD) study of 7,747 adults with a history of CHD and exposure to *C pneumoniae*, a 3–month course of azithromycin had no effect on recurrent CHD over the following 14 months.[347] Similarly, recent results from two large trials: the Azithromycin and Coronary Events Study (ACES) and the Pravastatin or Atorvastatin Evaluation and Infection Therapy trial (PROVE–IT), presented in August 2004 at the European Society of Cardiology annual conference in Munich, showed no beneficial effects on CHD risk following treatment with antibiotics among patients with prior cardiovascular disease.[348;349] Taken in conjunction with the results from the WIZARD study, these findings indicate that antibiotics should not be taken to prevent coronary heart disease.

2.5.4 Thrombotic and fibrinolytic factors

Following a plaque rupture, platelets bind to the site of the wound (and to each other) to form a blood clot (thrombosis), reducing the supply of blood and oxygen to the heart and causing damage or death (infarction) of the affected tissues. Most major heart attacks and deaths from coronary heart disease are caused in this way. The process through

which blood clots are naturally broken down in the body is known as fibrinolysis, the effectiveness of which depends both of the level of platelet activity and the body's own level of fibrinolytic activity. Epidemiological studies have shown that platelet aggregation is positively associated with CHD risk.[350;351] Furthermore, several fibrinolytic factors (factors involved in the process of fibrinolysis) have been proposed as being potentially relevant to CHD risk. High levels of fibrinogen and factor VII, two of the most important proteins used in the clotting process, have been shown to be associated with an increased risk of CHD.[343;352–355] In addition, high levels of D–dimer (a protein that is released into the blood stream during the process of fibrinolysis) have also been found to be associated increased CHD risk. D–dimer circulates for several days in the blood stream after a thrombotic event, and can therefore provide a valuable marker of the presence of unwanted thrombotic events, and potentially, therefore, of future coronary events. A review of six studies of the prospective relationship between D–dimer and CHD risk estimated that for individuals in the top third of the D–dimer distribution relative to individuals in the bottom third, the relative risk of CHD was 1.7 (95% CI 1.3 to 2.2).[356] Tissue plasminogen activator (t–Pa), the substance that activates the enzyme plasmin which actually dissolves the blood clot, may also be a potentially relevant risk factor,[357] as might von Willebrand factor, a substance released from endothelial cells which is involved in platelet adhesion of the damaged vessel wall and also in platelet aggregation.[358]

Randomised controlled trials have demonstrated that by inhibiting the body's ability to form a blood clot, CHD risk may be substantially reduced. In 2002, the Antithrombotic Trialists Collaboration of 287 studies involving over 135,000 patients in comparisons of antiplatelet therapy versus control reported that antiplatelet drugs (such as aspirin) reduce the risk of CHD by around one quarter and the risk of non fatal myocardial infarction by about one third.[359] Randomised controlled trials have also demonstrated the benefits of fibrinolytic therapy ("clot busting drugs") in patients with suspected acute myocardial infarction.[360] In individuals without clicical evidence of CHD however, randomised controlled trials have yet to establish whether targeting fibrinolytic factors could have an influence on preventing the occurrence of CHD. The role of fibrinolytic factors in terms of CHD prevention is therefore currently unclear.

2.5.5 Psychosocial factors and stress

Psychosocial factors include both environmental stressors and individual personality patterns or psychological reactions to stress. Typical environmental stressors are both stressors during everyday life and stressful work environments, the latter usually being characterised by both high demand and time pressure, and low control or decision making powers. This pattern is typically found in low status jobs, and is thought to account for some of the socio–economic gradient in CHD (see section 2.8.1 for description of social inequalities in CHD). The Whitehall II study[361] was one of the first studies to explore these issues after it was discovered that the large differences in CHD risk between men in the highest employment grade (administrators) and men in the lowest (messengers) observed in the Whitehall I study could only partially be explained by the established coronary risk factors.[362] The Whitehall II study thus set out to assess the possible role that "job related factors" including social isolation and support, coping styles, hostility and stress may have in the causation of coronary heart disease. Results have indicated that differences in psychosocial work environment[363] and in particular job control[364] may be independent risk factors for CHD, and that much of the observed differences in CHD risk between individuals at different grades of employment may be explained by these factors (though it is unclear to what extent these findings may simply be due to the strong correlation between job control and employment grade). Overworking,[365] depression,[366;367] social support[368] and hostility[369;370] have also been linked with increased risks of CHD, though mechanisms and preventive strategies remain unclear.

Possible mechanisms

One way in which psychosocial factors may influence CHD risk is through the neuroendocrine system; both work stress[371;372] and social isolation[373] have been shown to be associated with increased fibrinogen levels. In addition, depression[374] and anger[375] may be associated with low heart rate variability or sustained elevated heart rate. Can the potential effects of psychosocial factors be reversed? At the individual level, counselling and stress techniques have had some effect in reducing CHD risk,[376] however these findings have not been entirely consistent. It is therefore currently unclear as to what role these factors may have in terms of CHD prevention.

2.5.6 Fetal origins (the Barker hypothesis) and "early life" factors

The fetal origins or "Barker" hypothesis is perhaps the most controversial hypothesis relating to coronary heart disease. It states that adaptations made by the fetus in response to undernutrition lead to long–term changes in the metabolism and organ structure which predispose the individual to atherosclerosis and cardiovascular disease in later life.[377–379] The hypothesis was developed from observations made in studies of British men born in Sheffield during 1907-25[380] and Hertfordshire during 1911-30,[381] both of which found that birth weight was inversely associated with the risk of death from coronary heart disease. Further large studies in other populations, including Sweden[382] and the United States,[383] confirmed that these inverse associations were apparent in women as well as men. If the Barker hypothesis is true, then it is possible that the mechanism through which fetal growth influences subsequent CHD risk is through its effects on the established coronary risk factors in adulthood; slow fetal growth has been found to be associated with increased blood pressure,[132–135] increased serum total cholesterol,[135;384–386] impaired glucose tolerance and type II diabetes,[387] and impaired fibrinolytic activity[388] during adulthood. However, a recent systematic review has suggested that the association between birth weight and adult blood pressure reflects nothing more than a tendency to emphasise results that show the largest effects (which tend to be the smaller studies), and that really there is little or no effect of birth weight on adult blood pressure.[136] Similarly, a systematic review of the relationship between birth weight and blood cholesterol levels revealed only a very weak relationship.[389] In any case, since it is not yet clear which, if any, nutritional factors may play a role in determining CHD risk during adulthood, it is likely to be some time before specific advice given to women during pregnancy can be improved. In addition to the possible role that fetal factors may play in the causation of CHD in adulthood, numerous studies,[169;362;377;390–394] though not all,[395] have displayed an inverse relation between adult height and CHD mortality. This is consistent with the hypothesis that adverse environmental factors in early life have a direct effect on the risk of coronary heart disease.[377;378]

2.5.7 Genetic factors

In 1994, a study of over 21,000 Swedish twins born between 1886 and 1925 reported that after adjustment for other risk factors, individuals whose twin had died from CHD

before the age of 55 had substantially greater risks of dying from CHD themselves than individuals whose twin had not died of CHD before the age of 55, and that these risks were greater for monozygotic (identical) twins than for dizygotic twins.[396] The authors concluded that at younger ages in particular, genetic factors influenced the risk of death from coronary heart disease in both men and women. In recent years, the potential for genetic research to make an important contribution to the understanding of the causes of coronary heart disease has been greatly increased by the completion of the human genome project,[397;398] as well as by the development of new technologies for genomic analysis. Genetic factors are thought by some to form a substantial component of an individual's risk of developing CHD.[112] However the observation that individuals who migrate from one country to another tend to acquire the disease risks of their chosen country rather than retaining the risks of their country of origin,[399] suggests that (at the population level) environmental factors play a more important role.

So do genes really matter and, if not, how can genetic research further our understanding of CHD? Genetic research has certainly played an important role in therapeutic drug development (e.g. the identification of HMG–CoA reductase as the enzyme which regulates cholesterol production in the liver), and it has a large potential for helping to understand and predict variations in responses to drug treatments (e.g. the variable responses different individuals have to blood pressure lowering drugs). Furthermore, genetic research has already identified several important genetic mutations that lead to increased CHD risk. Probably the most established of these is familial hypercholesterolaemia, which is caused by codominant mutations in the LDL receptor gene, and which affects approximately 1 in 500 individuals in Western societies. Individuals with familial hypercholesterolaemia present with cholesterol concentrations approximately twice the average for their age and sex and, as such, can be easy to identify. However, there are other genetic variations affecting blood cholesterol concentrations that are, as yet, harder to identify. These include variations in the apolipoprotein B gene and the apolipoprotein E gene. In addition to their effects on serum cholesterol, it is also thought that genetic variations may also be important in determining the levels and effects of other established risk factors for CHD, including blood pressure, diabetes and even cigarette smoking.[400–402]

Mendelian Randomisation

Genetic association studies are currently undergoing a renaissance under the banner of "Mendelian randomisation", a term that refers to the suggestion that population–based studies of genotype–disease associations share with randomised controlled trials the advantage that confounding may be excluded as an explanation for any relations.[403] This is because at the time of gamete formation, the assortment of alleles is, by Mendel's second law, theoretically random. Therefore, the allocation of alleles should be independent of any environmental factors (both known and unknown) as well as any other variants at unlinked genetic loci, and hence any observed differences in disease risk should not be due to confounding (effectively these are naturally occurring randomised controlled trials). In addition, for genes known to modulate the effects of environmental exposure, genetic variants with known functional effects (e.g. the effect of codominant mutations in the LDL receptor gene on total cholesterol level) may also be considered as markers of altered environmental exposure, and hence (through the same Mendelian randomisation argument) may be used to assess an environmental risk factor's causal importance. This concept has been applied to several of the novel risk factors, notably homocysteine where a particular mutation in the MTHFR gene (that increases homocysteine level by approximately 20%) is known to exist.[307] In an analysis of 72 studies in which the prevalence of this mutation in the MTHFR gene was determined, individuals homozygous for the mutant allele (TT) had a 2.7 μmol/l higher serum homocysteine and a 21% higher risk of CHD than individuals homozygous for the wide type allele (CC), with no differences in any of the established coronary risk factors observed between TT and CC individuals.[301] Equivalent evidence for the causal involvement of fibrinogen and C–reactive protein appears less strong however.[404] Though the development of Mendelian randomisation provides an exciting promise for observational studies of gene–disease associations, several cautionary notes should be made. In particular, the size of the study is crucially important, as are the assumptions that the function of the gene is known and that alleles at nearby loci will not be preferentially associated with the alleles of interest. Furthermore, it should be noted that important gene–environment interactions may still exist even if gene–disease relationships appear not to.[405]

2.6 Contribution of different factors to CHD risk

Of the CHD risk factors described in sections 2.4 and 2.5, high blood cholesterol, high blood pressure and cigarette smoking have historically been regarded as the most important in public health terms because of the large causal relative risks associated with them, and because of their widespread prevalence in Western populations (cholesterol levels, in particular, are universally high throughout most Western populations). Given current trends in sedentary behaviour and obesity, it is likely that in the future, physical inactivity, obesity and diabetes will also become increasingly important causes of CHD at the population level. However, for many years it has been claimed that these established coronary risk factors, and in particular the three strongest causally related factors (high serum total cholesterol, high blood pressure and cigarette smoking), account for, at most, around half of all CHD cases.[30–40] To state that only half of CHD cases can be attributed to the established coronary risk factors has important implications. First, it suggests that the scope for prevention of CHD by reducing exposure to these factors is somewhat limited, reducing CHD at the population level by 50% at most, and second, it suggests that other widespread risk factors, including those as yet unknown, may be of critical importance.

While the origin of the "50% claim" goes back to at least 1975,[30] it doesn't seem to have been based on any empirical data.[42] Subsequent inappropriate analysis and/or interpretation of published reports, including the assessment of the impact of very high–risk groups on CHD (rather than the impact of above optimal risk factor levels on CHD)[31;35;40] and the inappropriate citation of studies that examine the extent that established risk factors explain psychosocial or socioeconomic gradients in CHD[33] (quite a different issue), has helped to perpetuate the myth. Furthermore, there is the potential for studies that seek to explain changing CHD death rates, and in particular the proportion explained by changing risk factors, to be misinterpreted, and reported as "evidence supporting the 50% claim". Increasingly however, various authors have provided evidence suggesting that the 50% figure may be an underestimate, and that the true contribution of the established risk factors to the population attributable risk of CHD is higher.[41–47] A summary of this evidence is now provided.

2.6.1 A re–assessment of the "50% claim"

In 1993, an analysis of the relationship between major risk factors and long term mortality from CHD in several large American cohorts, including the Multiple Risk Factor Intervention Trial (MRFIT) and the Chicago Heart Association (CHA) Detection Project in Industry study, found that compared with observed levels of serum and dietary cholesterol, blood pressure and cigarette smoking, the long–term CHD risks of those with "favourable" levels of these factors were between 60 and 89% lower, and that these individuals lived for up to nine years longer on average.[41] A secondary report from the same investigators presented six years later found that among 270,000 initially disease free men aged 40–57 years in MRFIT, the 16–year CHD death rate observed in non–smoking individuals with a serum total cholesterol < 5.17 mmol/L and a blood pressure no higher than 120/80 was 78% lower than in the rest of the sample.[44] In CHA, the 22–year CHD death rate in nearly 14,000 middle–aged men and women who were initially free from CHD was 77% lower (in men), and 79% lower (in women), for "low–risk" individuals compared with all remaining individuals.[44] These risk differences were even larger for men and women who were aged <40 years at time of recruitment. Similar sized differences between "low" and "high risk" individuals have also been found in other studies including the National Co–operative Pooling Project (where CHD risk was estimated to be 70% lower for men in the lowest quintile of risk, compared with all other men),[406] the Nurses Health Study (where risk was 82% lower in women who were "low–risk" according to a score derived from smoking, physical activity, obesity and diet),[45] and the Framingham study.[407] In the United Kingdom, the Whitehall I study of over 17,000 middle-aged British civil servants found that if the average CHD mortality rate in the whole population could have been reduced to that experienced by individuals who had never smoked cigarettes and who were in the lowest quintiles of blood cholesterol and blood pressure levels, then about two-thirds of the CHD deaths would have been avoided.[42] Most recently, two studies estimating the prevalence of established risk factors in a large number of individuals with coronary heart disease found that between 80 and 90% of such individuals were exposed to at least one of the main risk factors (high blood cholesterol, high blood pressure, smoking or diabetes),[46;47] in stark contrast to the "only 50% claim". All of these studies provide evidence against the "50% claim". However, none of these studies were able to take into account "within–person" variation in coronary risk factors, the effects of which are now described.

The use of baseline measures in analyses of incident disease

Analyses of incident disease in epidemiological studies typically use "baseline" assessments of individuals as estimates of the individual's "exposure" to that particular risk factor, in order to separate "cause" from "effect". However, baseline measurements often do not reflect an individual's true usual level over time (because of measurement errors, short term "random" fluctuations from the individual's average level and longer term systematic changes). Though these effects are random, meaning that the baseline measurement is just as likely to over- as under-estimate the subject's true level, the differences between individuals estimated from a baseline sample (the between person variation) tend to exaggerate the true differences that really exist between those patients over a period of time. In other words, the differences in the level of the risk exposure between the study participants are not as large as one would estimate from the baseline sample alone. In consequence, for single continuous risk factors that display linear "dose–response" relationships with disease risk (as total cholesterol and blood pressure do for CHD), the estimated association derived from baseline measures underestimates the true association,[15–18] a phenomenon that has become known as "regression dilution bias"[17;19;20] (see section 3.5.2 for further details). These effects are important when assessing the true combined importance of the established coronary risk factors, and while the studies previously described provide evidence against the "only 50% claim", none of them were able to take this source of bias into account.

2.7 Strategies for CHD Prevention

When considering the prevention of coronary heart disease, it is usual to draw the distinction between primary and secondary prevention.[48] Primary prevention refers to the prevention of CHD amongst individuals without CHD. Secondary prevention on the other hand refers to the prevention of new CHD events in individuals who already have clinically manifest CHD (e.g. angina or non–fatal myocardial infarction). In secondary prevention, it is generally recognised that all individuals should be offered advice to modify their lifestyle or given drugs to reduce their CHD risk. In the United Kingdom, for instance, the current National Service Framework (NSF) for coronary heart disease recommends that all individuals with established CHD should receive statin treatment providing that

their total cholesterol is greater than 5.0 mmol/L or their LDL cholesterol is greater than 3.0 mmol/L[408] (though it is argued by many that statins should be given to these individuals irrespective of their cholesterol concentrations).[71;73] In primary prevention, where the absolute risk of CHD is lower, there are two accepted approaches to prevention: the "high–risk" and "population" approach, as illustrated in figure 2.5. High-risk approaches to prevention aim to identify and treat those individuals thought to be at greatest risk of developing disease. The key processes governing the potential effectiveness of the high–risk approach are: (1) the ability to identify those at greatest risk of CHD; (2) the choice of the "threshold" level of predicted risk deemed sufficient to warrant treatment; and (3) the effectiveness of the risk reducing measures available. In contrast, the population approach to prevention seeks to bring about a downwards shift in the population distributions of the most important risk factors (see figure 2.5). Each of these approaches to prevention is now considered in more detail.

2.7.1 High–risk approaches to prevention

Identifying the high–risk group

Originally, high–risk approaches to the primary prevention of coronary heart disease have been based on the identification and treatment of individuals with elevated levels of single risk factors, for example elevated levels of total cholesterol or blood pressure. More recently however, high–risk approaches to prevention have been based on the estimated overall risk of CHD, taking into account several risk factors simultaneously. The concept of being able to develop a tool for predicting the absolute CHD risk in any specific individual was one of the primary objectives of the Framingham Heart Study, with the first "risk equations" from this cohort being published in 1976.[409] These equations were superseded in 1991 in order to take into account more data on individuals over 60 years old (including data from the Framingham Offspring Study) as well as to include new risk factors such as HDL cholesterol.[410] Six new equations were derived for predicting the risk of: (1) myocardial infarction; (2) coronary death; (3) any CHD event (including silent myocardial infarction and coronary insufficiency); (4) stroke; (5) cardiovascular death; and (6) any cardiovascular event. While several other risk scoring methods have been proposed,[411–413] it is the equations from Framingham that are the most widely used. In the United Kingdom, for instance, current guidelines on the primary prevention of CHD are

based on the Framingham equation relating to "any CHD event". The National Service Framework for coronary heart disease (published in 2000) states that individuals without symptoms of CHD, but with a (Framingham) predicted risk of at least 30% of developing it within 10 years, should be identified and considered for treatment with statins.[408] Similarly, other European, American and Canadian guidelines have based their risk prediction methods on equations derived from the Framingham study.[4;414–416] Recently however, new European guidelines[417] have recommended risk prediction based on the newly published SCORE equations,[418] which are derived from many different European studies and which take into account the differing background risks of different European populations. The decision to replace the use of the Framingham equations with a set of new ones may have been partially motivated by the growing evidence that the Framingham equations tended to overestimate absolute coronary risk in European populations.[419–424]

Potential effectiveness

Assuming that individuals can be suitably "ranked" in order of CHD risk, the effectiveness of the high–risk approach to prevention depends on the threshold at which treatment is provided and the risk reducing capabilities of the interventions available. As already stated, the NSF for CHD in the United Kingdom recommends a ten–year CHD risk threshold of 30% or more for treatment intervention. In the British Regional Heart Study, this criterion identifies (at baseline) only approximately 6% of the healthy men as being "high–risk", thereby limiting the potential for this approach to substantially influence population levels of CHD. European guidelines have generally set lower thresholds. Prior to the recent recommendations of the Third European Joint Task Force Report on Coronary Prevention, European guidelines recommended treatment based on a Framingham ten-year CHD event risk of 20% or more.[4] Regardless of the threshold level chosen however, the potential for reducing the CHD risks of those at greatest risk has been greatly increased in recent years by the availability of several safe, effective and apparently independent treatments.[51] It has been suggested that by combining several drugs (such as statins, aspirin and one or more blood pressure lowering drugs), the risk of CHD in individuals could be reduced by at least 70%,[51] and possibly by even more.[425]

2.7.2 Population approaches to prevention

The high–risk approach to CHD prevention is a natural approach for medical practitioners who are concerned with the occurrence of disease in individuals, providing them with a quantitative tool that enables them to "rank" individuals in order of priority so that resources can be allocated appropriately. However, the occurrence of coronary heart disease is not confined only to those at greatest risk. Indeed, the majority of CHD cases occur amongst individuals whose absolute excess risk is small,[49] and whose risk exposure levels are "normal" or "usual" for their own population, meaning that they are near to the average level in that population. However, when compared with a different population their exposure may be considered anything but normal.

A striking example of this is provided by the population levels of serum total cholesterol levels in Finland and Japan (as measured in the Seven Countries study),[29] where the distributions barely overlapped (mean levels were approximately 6.8 and 4.4 mmol/l respectively). These kinds of "between–population" differences have led some to advocate that (primary) prevention policies should be "population" rather than "individual" based.[48] Population approaches to prevention aim to cause a downwards "shift" in the distribution of causally-related risk factors throughout the entire population (see Figure 2.5), rather than targeting specific individuals. For CHD, this would logically involve reducing the population average levels of blood cholesterol and blood pressure, and reducing cigarette smoking rates. However, though the absolute benefits experienced by the population following such reductions may be great, the marginal benefits experienced by each of its individuals due to "population shifts" would be relatively small (except for cigarette smokers who give up). Individuals may not be prepared to alter their lifestyle so that the population as a whole can enjoy lower rates of CHD.

Population approaches are therefore often implemented through government interventions such as increased taxes on cigarettes and tighter regulation of the salt and fat content in processed foods. These types of approaches to causing changes in the distribution of a risk factor within a population are feasible; numerous studies have demonstrated that community levels of serum total cholesterol and blood pressure can be reduced (in some cases substantially) through changes in diet.[104;426–430] In North Karelia, Finland, cholesterol levels have been reduced by 18% in both men and women and blood pressure by 8% in men and 13% in women since the introduction of community based cardiovascular

disease prevention project in 1972.[426] Similarly, in a study of adults living on the island of Mauritius, mean total cholesterol fell from 5.5 mmol/l to 4.7 mmol/l over 5 years following the implementation of a population–wide intervention programme aimed at the promotion of a healthy lifestyle,[427] though the fall in cholesterol was probably attributable to a change in the island's supply of cooking oil from palm to soy bean oil rather than any behavioural effect of health promotion.[431] Meta–analysis of metabolic ward studies in Western countries has indicated that a reduction in total cholesterol of this size could be achieved in the United Kingdom if 60% of saturated fats could be replaced by other fats and 60% of dietary cholesterol could be avoided.[104] Similarly, analysis of clinical trials of salt reduction indicate that simple dietary changes (reducing salt intake by approximately 50 mmol (3 grams) per day) could easily reduce systolic blood pressure in the population by 5 mmHg, while reduction also in the amount of salt added to processed foods would lower population blood pressure by twice as much.[127] These estimates are consistent with observed reductions in population blood pressure in a community trial of salt reduction in Portugal.[430] Current public health recommendations in the UK[432] and US[433] are to reduce salt intake from 9 to 12 grams a day to 6 grams or less, however it has been suggested that, given the consistency in the dose–response relationship observed between salt intake and blood pressure, the long–term target for population salt intake should be no more than 3 grams a day.[434] This may best be achieved through governmental regulation of the salt content in processed foods, a policy that has been shown to be a particularly cost–effective method of preventing cardiovascular disease.[435] Furthermore, such changes to processed foods are unlikely to be resisted, as there is evidence that small and repeated decreases in salt intake are not discernible on the grounds of taste.[436]

The key feature of the population approach to prevention is that it seeks to remove or reduce the underlying causes that make the disease common in the population. The approach therefore has a large potential for the population as a whole. However, in order to estimate the effect that population approaches to CHD prevention may have in the long–term, an accurate assessment of the true relationships between risk factors and CHD risk is required.

2.8 Social and geographic inequalities in CHD

Inequalities in the incidence of coronary heart disease in the United Kingdom have been observed for many years. In this section, social and geographic inequalities in CHD risk are described, with particular emphasis on why they occur, how large they are, and how they have influenced prevention strategies in the United Kingdom.

2.8.1 Socioeconomic inequalties

What is socioeconomic status?

In order to be able to interpret the meaning of social inequalities in CHD, one needs to first define what is meant by "socioeconomic position" or "socioeconomic status" (SES). The term "socioeconomic status" usually comprises of a variety of measures including level of education, income, occupation and living conditions. However, precisely what is meant by socioeconomic status has long been debated in the history of sociological theory, with approaches generally reflecting the orientation of either Weber[437] or Marx.[438] The Weberian approach of sociology sees social stratification as being organised around the three independent entities of economic interest, status and power. Under such an approach, subjective measures of status as well as more objective measures such as income and education will form the basis for a summary socioeconomic measure. Indeed many studies have emphasised the role of status (as recorded by job occupation) thereby preserving this historical basis. In comparison, the Marxist approach views socioeconomic status as the opposing interests of those who differ with respect to ownership of the means of production. To (over)simplify, individuals who differ in their roles in the production process are thought to be locked in conflict with each other. Though a large number of indices are available that combine these different aspects of SES,[439;440] these summary measures can often result in difficulties over and above those of their individual components. Indeed, several reviews of SES and health have recommended that summary measures of socioeconomic status should not be used.[439–441] Rather, epidemiological studies generally use single measures of SES as simple and convenient markers of the underlying socioeconomic factors. In this thesis, occupational social class (determined in the UK by the Registrar General's six-category classification)[442] is used as a simple measure of an individual's position within the "socioeconomic spectrum".

Reasons for social differences in CHD

Social inequalities in the incidence of coronary heart disease and stroke, with higher rates amongst lower SES groups, are well documented.[34;362;443–445] However, in the United Kingdom, though absolute CHD rates have fallen during the last 20 years,[446;447] the fall has been concentrated among higher social class groups so that the relative differences between those at the top and those at the bottom of the social scale have widened.[54] This observation has led to considerable emphasis being placed in recent public health policies on reducing social class inequalities in CHD.[54;55]

But why do these inequalities occur? Many studies have estimated the extent to which social inequalities in the development of atherosclerosis and CHD may be attributed to differences in the established risk factors, most of which have found this to provide only a partial explanation.[362;448–451] Of the individual risk factors, cigarette smoking has often been found to provide the greatest explanation for social class inequalities in CHD, because it is both strongly correlated with social class (negatively) and because it is an important determinant of CHD risk.[153;450;451] However, the failure of the established risk factors to entirely explain social class inequalities in CHD has prompted several authors to claim that other risk factors must be of critical importance in the causation of CHD.[443] The Whitehall II study was designed with this question in mind, and in particular to explore the potential for psychosocial factors to act as mediators in the social class–CHD relationship[361] (see section 2.5.5 for discussion of psychosocial factors and CHD risk). A further explanation for social inequalities in cardiovascular disease in adulthood may lie in the role of childhood socioeconomic circumstances. Numerous studies,[169;362;377;390–394] though not all,[395] have found an inverse relationship between adult height and cardiovascular mortality. This is consistent with the hypothesis that adverse environmental factors in early life may have a direct effect on the risk of cardiovascular disease in adulthood,[377;378;452–454] (though, as demonstrated by two reports from the same study reaching opposite conclusions,[455;456] this evidence is not entirely consistent).

2.8.2 Geographic inequalities

Differences between countries

Large international differences in the occurrence of CHD and stroke have been observed for many years. The Seven Countries study was one of the first studies to explore the reasons for these differences, with results indicating that many of the differences in CHD rates were likely to be due to differences in national diet, in particular the percentage of total energy intake derived from saturated fat. However, despite these findings, large international differences in the risk of CHD are still evident today. Recent data from the World Health Organization, for instance, revealed that the age standardised mortality rate of CHD and CVD in the Russian Federation was six times that in France.[457] The same data showed how eastern European countries such as the Ukraine, the Russian Federation, Hungary, and the Czech Republic currently have among the highest CVD rates in the World and that, in contrast to most economically stable European countries, the rates of CHD in eastern European countries are increasing so that the differences between these countries and other European countries are likely to increase. Many of the differences in CHD and stroke rates between different European countries can be explained by differences in the classical risk factors,[8] though it has been hypothesised that the French, who have very low CHD rates, may experience additional benefits due to a high consumption of wine.[458] [i] In comparison with European and American populations, CHD rates in Japan have traditionally been very low, probably due to low levels of serum total cholesterol resulting from a national diet low in saturated fat and cholesterol. Japan differs from most western countries in that stroke rates are higher than CHD mortality rates.[460] This is thought to be due to a combination of low total cholesterol with high population levels of blood pressure and high cigarette smoking rates, the two main risk factors for stroke.

Differences within countries

In addition to differences in CHD between countries, many countries experience substantial risk differentials (albeit smaller ones) within their own population. In the United Kingdom

[i] The low CHD rates in France are despite high smoking rates and a high population average blood cholesterol (the "French Paradox"). While this may partially be explained by a high consumption of wine, it may also reflect poor death certification in France or possibly the observation that cholesterol levels have only been high for a relatively short period of time (the "time–lag" effect).[459] However since trial evidence strongly suggests that CHD risk is responsive to changes in blood cholesterol over a relatively short period of time, the latter explanation is becoming less and less likely.

for instance, it is well established that the risk of CHD is higher in the North of England and Scotland than in the South of England[461] (as illustrated in Figure 2.6). Similarly, the risk of CHD is higher in Southern and Eastern states of America than Northern and Western ones,[462] higher in urban India compared with rural India,[463] higher in northern China (Beijing) compared with southern China (Shanghai and Guangzhou),[464] and higher in rural Australia compared with urban Australia.[465] Various studies have demonstrated that geographic differences in CHD within a country are substantially, though not entirely, explained by geographic differences in the established risk factors.[461;466–468]

2.9 Conclusions

Observational and experimental studies have emphatically demonstrated the crucial role that blood lipids, blood pressure and cigarette smoking have in determining CHD risk, both in explaining the fundamental reasons for between population differences in CHD as well as predicting the risk of CHD within a population. Other risk factors for CHD including physical activity, obesity and diabetes have also been shown to greatly influence CHD risk in individuals, and though the consistency of findings regarding the protective effects of alcohol consumption strongly indicate a causal effect, the protective mechanisms are not yet entirely understood. There is a widespread belief however, that these risk factors only partially explain CHD risk. As a result, a vast number of new hypotheses have been suggested as playing important roles in CHD risk, though the evidence in favour of these hypotheses is variable. Despite some research indicating that the role of the established risk factors may have been underplayed, the effects of using baseline assessments of individuals in analyses (which leads to biased estimates of risk associations), has not been adequately addressed with regard to assessing their combined "importance". Furthermore, strategies to prevent the occurrence of CHD in the population have not been formally compared in terms of assessing the impact that these sources of bias may have on the estimated effectiveness, or in taking account of advances in preventive therapies for CHD. Additionally, though inequalities in the incidence of CHD are recognised, several issues regarding their causes and their "real" importance remain unanswered. Important themes therefore need to be re–addressed: (1) what impact does within–person variation in the major risk factors have on their estimated relationships with CHD risk?; (2) what

impact does this have on the estimated combined contribution of the major risk factors to population levels of disease, and what implications does this have for CHD prevention?; and (3) do these influences affect the estimated size of social class inequalities in CHD or the extent to which they may be attributed to differences in the established risk factors?

Table 2.1: Primary findings of the six "statin" trials

Trial	Study population	Treatment	Primary endpoint	Risk reduction
4S (1994)	4444 men and women aged 35-70 with history of CHD and TC 5.5-8.0 mmol/L	Simvastatin	Total mortality	30% (15%,42%)
WOSCOPS (1995)	6569 men aged 45-64 with no previous MI, average TC 7.0 mmol/L	Pravastatin	CHD event including silent MI	31% (17%,43%)
CARE (1996)	4159 men and women aged 21-75 with recent MI, TC < 6.2 mmol/L	Pravastatin	CHD event	24% (9%,36%)
LIPID (1998)	9014 men and women aged 31-75 with recent MI or unstable angina, TC 4.0-7.0 mmol/L	Pravastatin	CHD death	22% (13%,31%)
AFCAPS/Tex CAPS (1998)	6605 men and women aged 45-73 with no clinical CVD, average TC 5.7 mmol/L	Lovastatin	CHD event or unstable angina	37% (15%,48%)
HPS (2002)	20536 men and women aged 40-80 with diagnosed CHD, other arterial disease, or diabetes, TC > 3.5 mmol/L	Simvastatin	Total mortality	13% (6%,19%)

MI = myocardial infarction, TC = total cholesterol, CHD event = Coronary death or non fatal MI, RR = relative risk, CI = confidence interval, 4S = Scandinavian Simvastatin Survival Study, WOSCOPS = West of Scotland Coronary Prevention Study, CARE = Cholesterol and Recurrent Events Trial, LIPID = Long term Intervention with Pravastatin in Ischaemic Disease study, AFCAPS/TexCAPS = Air Force/Texas Coronary Atherosclerosis Prevention Study, HPS = Heart Protection Study

Figure 2.1: Incidence of ischaemic heart disease, age adjusted with 95% confidence intervals, according to fifths of distribution of serum cholesterol concentration in the Multiple Risk Factor Intervention Trial. Figure amended and reproduced with permission from the BMJ Publishing Group (*BMJ* 2002 **324:** pp 1570 – 1576).

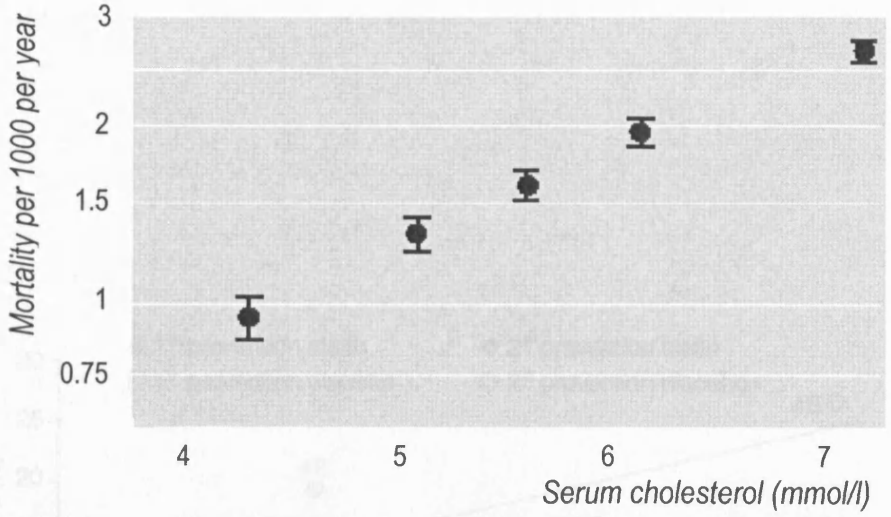


Figure 2.2: Relationship between the CHD event rates in the five statin trials published before 1999 and the LDL cholesterol levels achieved in those studies (4S = Scandinavian Simvastatin Survival Study; AFCAPS = Air Force/Texas Coronary Atherosclerosis Prevention Study; CARE = Cholesterol and Recurrent Events trial; LIPID = Long-term Intervention with Pravastatin in Ischaemic Disease study; WOSCOPS = West of Scotland Coronary Prevention Study). Figure reproduced from Ballantyne. *Am J Cardiol* 1998, **82:** pp 3Q–12Q

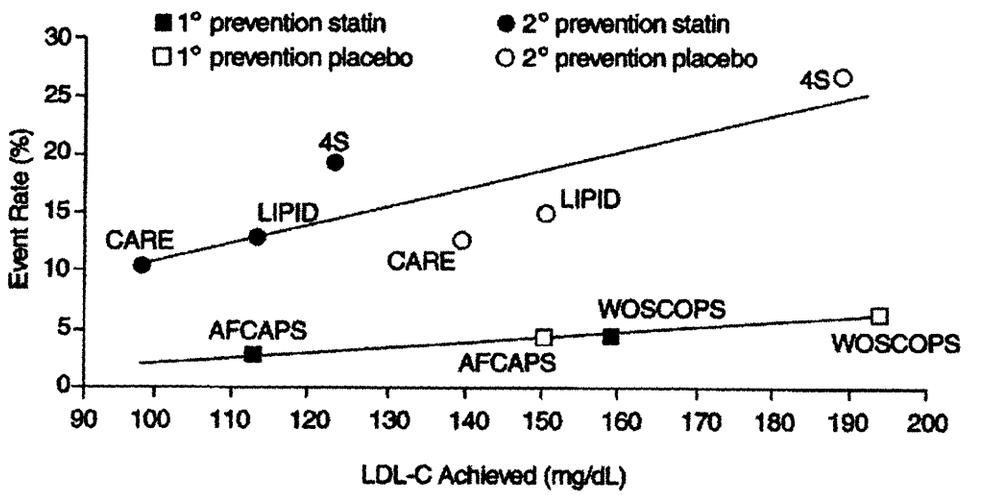


Figure 2.3: Ischaemic heart disease (IHD) mortality rate in each decade of age versus usual blood pressure at the start of that decade. Reprinted with permission from Elsevier (*The Lancet* 2002, **360:** pp 1903–1913).

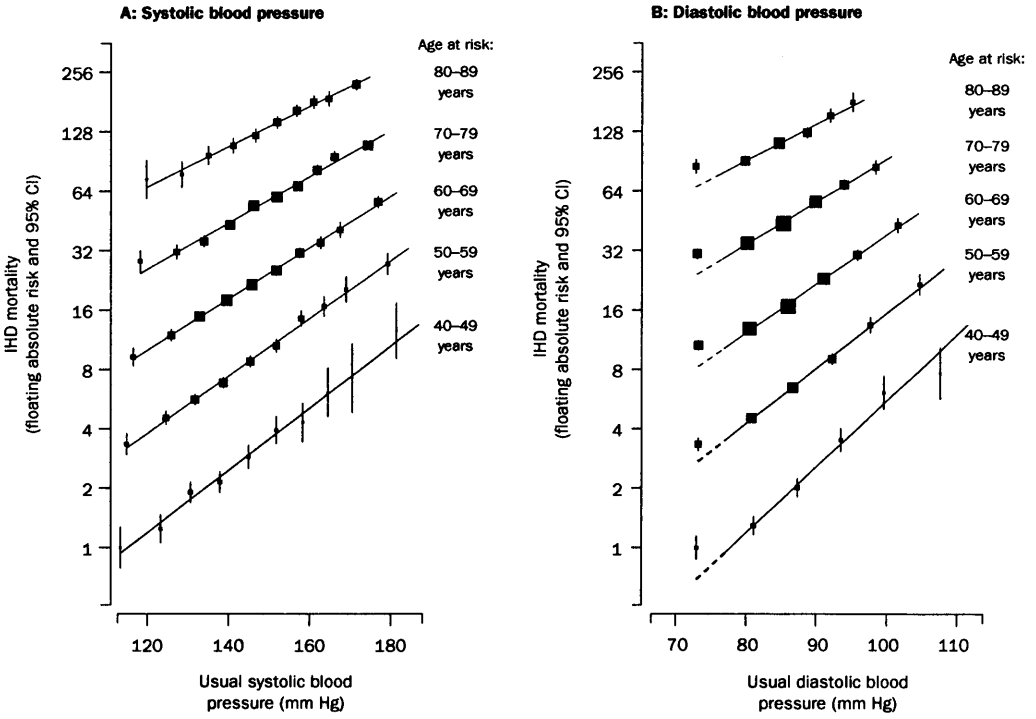


Figure 2.4: Relation between odds ratio for cardiovascular mortality, and corresponding differences in systolic blood pressure observed in 27 randomised controlled trials of blood pressure lowering drugs. Reprinted with permission from Elsevier (*The Lancet* 1997, **358:** pp 1305–1315).

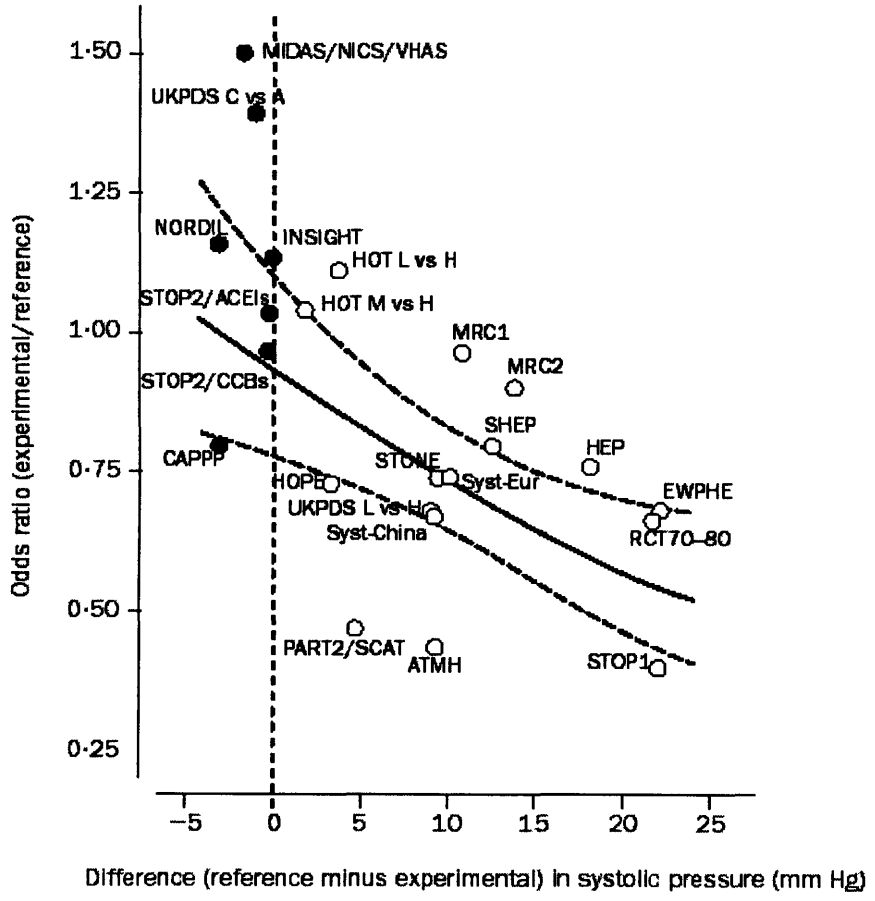


Figure 2.5: The high–risk approach (left) and the population approach (right) to the primary prevention of coronary heart disease

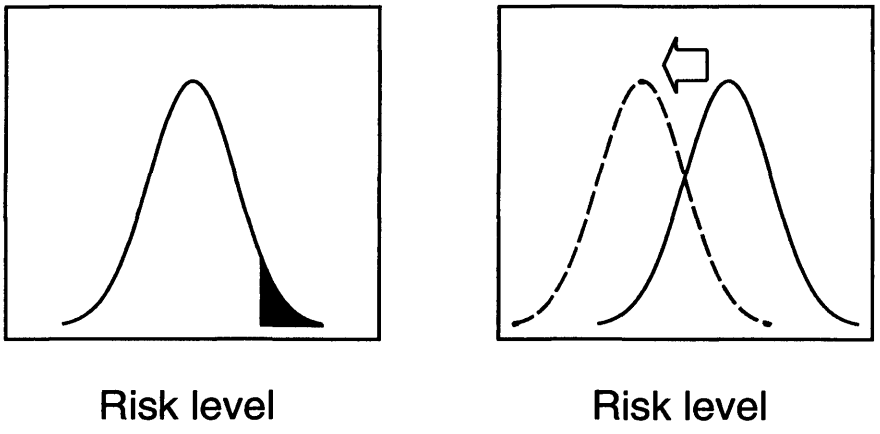
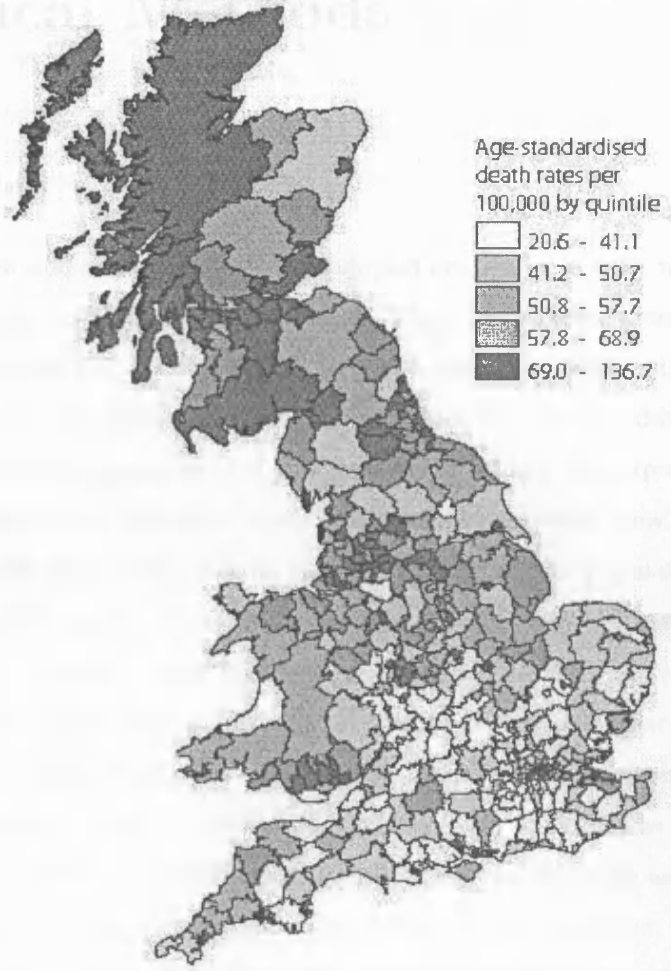


Figure 2.6: Age standardised CHD death rates per 100,000 population by local authority for men under 65, 1998/2000, United Kingdom. Figure reprinted from the 2003 edition of the British Heart Foundation Coronary Heart Disease Statistics database



Chapter 3

Statistical Methods

3.1 Summary

Numerous classical and regression based techniques are available with which the importance of a risk factor for CHD may be quantified. These methods are commonly employed to estimate the relative risk of CHD for one level of a risk factor relative to another. However, the relative risk of CHD (or other similar statistics such as the odds ratio or hazard ratio) does not allow the assessment of how removal of a high–risk exposure throughout a population would affect population levels of disease (for example, how removal of cigarette smoking would affect CHD). Under such circumstances, the population attributable risk fraction (PARF) may be of more relevance, as it takes into account the proportion of the population "exposed" to the high–risk factor as well as the relative risk of disease. For continuous risk factors that are linearly related to disease risk however, the use of a "threshold" level to define "high–risk" is arbitrary, and so it may be necessary to consider the PARF as a function that varies with the threshold used to define the high–risk group. Regardless, these measures of "aetiological force" depend on accurate assessments of average exposure levels in individuals. However, within–person variation in coronary risk factors can lead to underestimation (or in some cases overestimation) of the assessments of risk associations, when these estimates are derived from single "baseline" measurements of study participants. Methods to estimate and correct for this bias are available however provided that repeated measurements of study participants are available.

3.2 Introduction

In this chapter, the principles that underlie the key statistical analyses presented in this thesis are described. In particular, the effects of within–person variation in coronary risk factors on estimated associations with CHD risk are reviewed for both continuous risk factors (where the effects are usually referred to as "regression dilution bias") and categorical factors (where the term "misclassification bias" is often used to describe the effects). In sections 3.3 and 3.4 of this chapter, the relative risk and population attributable risk fraction are defined and the range of different regression methods used throughout this thesis are described. In sections 3.5 and 3.6, the effects of within–person variation in continuous and categorical risk factors are described and methods to estimate and adjust for these effects reviewed. In section 3.7, these effects are discussed in the more general context of estimating risk–relationships after taking into account several coronary risk factors subject to within–person variation. Specific methods relevant to the analyses in this thesis are described separately in each of the results chapters 5 to 9.

3.3 Measures of association

3.3.1 The relative risk, the odds ratio and the hazard ratio

The most common measure of "aetiological force" used in epidemiology is the relative risk (RR), defined as the risk (or probability) of disease in one group of individuals divided by the risk of disease in another. The relative risk provides a simple means of summarising the relative differences in risk between two groups of individuals and, for this reason, is regularly used to quantify the "importance" of the characteristic (or group of characteristics) that distinguish the first group from the second. However, in statistical analyses, it is more usual to consider the odds of disease in a particular group. This is defined as the number of *diseased* individuals in a group divided by the number of *disease–free* individuals in that group (rather than the total number as is used for the relative risk). The odds ratio (OR) is then defined as the odds of disease in one group divided by the odds of disease in another and is preferred in statistical analyses to the relative risk because it exhibits certain "desirable" properties (specifically the sampling distribution of its (natural) logarithm is normally distributed) that allow it to be "modelled" in (logistic) regression analyses. However, it is noted that when the disease is rare, the odds ratio and the rela-

tive risk do give similar estimates and it is perhaps for this reason that the two are often used interchangeably (usually the odds ratio is mistakenly referred to as the relative risk). The relative risk and odds ratio are not the only measures of relative differences between groups however. In *time–to–event* or *survival* analyses (see section 3.4.2), it is more usual to refer to a hazard ratio (*HR*), defined as the ratio of the "instantaneous hazard" in two groups. This measure takes into account the fact that disease events (failures) may occur at any point in time over a particular follow–up period, and that information regarding the "importance" of a particular risk exposure are gained, not only from whether these events occur or not, but also from when the events occur.

3.3.2 The population attributable risk fraction

For risk factors that are causally related to disease risk, the relative risk, odds ratio and hazard ratio **all** provide appropriate measures of its relative importance. However, they do not take into account the number of individuals in the population exposed to the risk factor of interest. Therefore, a high relative risk may not be indicative of a serious public health problem if the prevalence of the risk factor in the population is small, or if the disease is rare. In contrast, a causal factor with low relative risk may be important if a high proportion of the population is exposed to the factor and the disease is common. In this context, the population attributable risk fraction is a more useful measure of the public health importance of the risk factor since it takes into account both the proportion of the population exposed to the factor and the relative risk associated with the factor.

Levin's population attributable risk

The population attributable risk fraction (PARF) is the proportion of all disease events in the population that could be eliminated if the high risk factor was not present. It can be calculated from the prevalence of the risk factor in the population (π) and the relative risk of the exposed to the unexposed (RR) through Levin's equation[469]

$$\text{PARF} = \frac{\pi(RR-1)}{1+\pi(RR-1)} \tag{3.1}$$

If the risk factor is rare, so that π is close to zero, the PARF will be small even if the relative risk associated with the risk factor is large. This indicates that removal of the factor from the population would have only a small effect on population disease rates. When the prevalence of the risk factor is high however, the PARF will approach the value of the "relative risk reduction" $(1 - 1/RR)$, experienced by individuals without the risk factor compared with individuals with the factor.

Levin's measure of the population attributable risk fraction compares a low risk group with a single high risk group. Suppose however that the marginal contribution of several distinct high risk groups to the combined PARF was required. For instance, a high risk group containing all individuals with either 'high' total cholesterol or 'high' blood pressure could separated into three subgroups: those with high cholesterol only; those with high blood pressure only and those with both high cholesterol and high blood pressure. Through a simple extension to Levin's equation, the separate contributions of these three high risk groups to the overall combined PARF of high blood cholesterol and high blood pressure can be calculated. Supposing that there are N possible high–risk groups, the marginal PARF for the jth high risk group is calculated by

$$\text{PARF}_j = \frac{\pi_j(RR_j - 1)}{1 + \sum_{k=1}^{N} \pi_k(RR_k - 1)} \tag{3.2}$$

where RR_j is the relative risk for the jth high–risk group compared with the reference (low–risk) group and π_j is the proportion of the population falling into the jth high–risk group. The combined PARF is then simply calculated as the sum of the marginal contributions:

$$\text{PARF} = \frac{\sum_{k=1}^{N} \pi_k(RR_k - 1)}{1 + \sum_{k=1}^{N} \pi_k(RR_k - 1)} \tag{3.3}$$

3.3.3 The low–risk group and "PARF curves"

When estimating a population attributable risk fraction associated with one or more risk exposures, the choice of the low–risk (or reference) group is particularly important as it affects both of the measures used in its calculation: (1) the prevalence of the high–risk

factor(s); and (2) the relative risk of disease between the high risk group and the low risk group. For categorical risk factors, the choice of the low risk group may be clear (e.g non–cigarette smokers as the low–risk group for smoking exposure), and under such circumstances, the PARF may be easily calculated (either through classical methods or through fitting a binary term in a generalised linear model), and easily interpreted (it is the reduction in disease that would be expected if nobody smoked cigarettes). However, for continuous risk exposures, in the absence of a "threshold" level above which risk is known to be uniformly elevated and below which risk is known to be constant, it becomes necessary to separate the low and high–risk groups based on an arbitrary cut–off level. In fact, when no "threshold" exposure level is present (for example when there is a log–linear relationship between the risk exposure and the odds of disease), it is more appropriate to consider the PARF as a smooth function of the cutoff used to calculate it (hereon referred to as a "PARF–curve"), rather than as a single true value.

3.4 Regression methods

Regression is used in epidemiological studies for two reasons: (1) to estimate the relationship between one or more predictor variables and the risk of disease; and (2) to predict the probability of disease given the values of certain predictor variables. A regression model extends the boundaries achievable from univariate analyses by allowing multiple risk factors to be simultaneously related to the risk of disease, thus enabling the "independent" contribution of separate risk factors to be assessed. Furthermore, regression models can easily identify whether certain risk factors are more or less important depending on the level of other factors (tests for interaction) and can be used to determine the "best" model given a wide range of potential factors. In this section, a brief review of regression models including generalised linear models (of which linear regression is the simplest) and "survival" regression models is presented. A complete discussion of regression based techniques may be found elsewhere.[470–473]

3.4.1 Generalised linear models

Generalised linear models (GLMs) provide a method of simultaneously relating one or more predictor variables (which can be continuous and/or categorical) to a single outcome

variable. In the case of linear least squares regression (the simplest GLM), variables are related to a continuous normally distributed outcome (such as blood pressure) through the equation:

$$\mathbf{Y} \sim \mathbf{X}\beta \tag{3.4}$$

where $\mathbf{Y}$ is a vector of observed outcomes, $\mathbf{X}$ is a "design matrix" of observed predictor variables and β is the vector of regression coefficients that show the strength of relationship between the predictor variables and the outcome. When the outcome variable cannot be modelled as a linear function of the explanatory variables however (for example when the response data are not continuous), the generalised linear model assumes that some function of $\mathbf{Y}$, rather than $\mathbf{Y}$ itself, is linearly related to $\mathbf{X}\beta$. This function is known as the "link function" and provides a method of relating predictor variables to many different types of outcome data (not just normally distributed outcomes). For logistic regression, which allows predictor variables to be related to a single binary outcome, the usual (though not only) link function is the "logit" link defined as $l(p) = \log(p/(1-p))$, where p is the expected probability of the outcome. Logistic regression is commonly used in epidemiological studies where the outcome of interest is the occurrence of a particular disease. However, a further class of regression methods (survival regression methods) extends this approach by considering the time until an event occurs as well as the occurrence of the event itself. These "time–to–event" or "survival" analyses are now described.

3.4.2 Survival analyses

Survival analyses differ from conventional regression methods in that both the occurrence (or not) of an event and the time until that event are taken into account, and it is for this reason that they have become the preferred method of analysis in many prospective epidemiological studies. Individuals who do not experience an event after a certain period of follow–up has elapsed or individuals who have otherwise been lost to follow–up during the follow–up period are said to be "censored" at the time they were last observed to be free from disease. Denoting the time until the first event by T, the time until censoring by U, and the occurrence of an event by an indicator variable δ, the "survival function",

$S(t)$, and the hazard function, $h(t)$, are defined as follows:

$$S(t) = \Pr(T \geq t) \tag{3.5}$$

$$h(t) = \lim_{\mathrm{dt} \to 0} \frac{\Pr(T \in [t, t+\mathrm{dt}) \mid T \geq t, U \geq t)}{\mathrm{dt}} \tag{3.6}$$

Parametric and Non–parametric estimates of the survival function

Survival analysis regression methods tend to be either parametric or non–parametric in nature. Parametric methods impose distributional assumptions on the form of the survival function and the hazard function. One such assumption which is widely used in parametric survival regression methods is the proportional hazards assumption. This assumes that the survival function and hazard function can be represented in the form

$$S(t) = S_0(t)^{\gamma} \tag{3.7}$$

$$h(t) = \gamma h_0(t) \tag{3.8}$$

where γ is a function of the predictor variables $\mathbf{X}\beta$, and $S_0(t)$ and $h_0(t)$ are the baseline (average) survival and hazard functions respectively. The key to the proportional hazards assumption is that relative differences in risk (hazard) between individuals remain constant over the period of follow–up. In contrast, non–parametric methods make no assumptions about the family of distributions that $S(t)$ and $h(t)$ can belong to.

Kaplan–Meier survival curves and the log–rank test

The most common non–parametric estimate of the survival function is the Kaplan–Meier (or product–limit) estimate,[474] defined as the product of survival probabilities

$$\hat{S}_{KM}(t) = \prod_{t_i < t}^{k} \frac{r(t_i) - d(t_i)}{r(t_i)} \tag{3.9}$$

where t_i is the failure (event) time of the ith individual, $r(t_i)$ is the number of individuals "at risk" at t_i (after the failure of the ith individual), and $d(t_i)$ is the number of failures

(events) observed up to and including t_i. The Kaplan–Meier estimate is usually shown graphically as a step function with a drop occurring at each event, though often, it is one minus the Kaplan–Meier estimate that is displayed in order to show the cumulative events rather than the survival probabilities. Differences in survival function between two or more Kaplan–Meier curves are usually assessed through the log–rank test. This non–parametric test tests the null hypothesis that the groups have the same survival distribution by analyzing and comparing the number of observed and expected events for each group each time an event occurs. Details on how the log–rank test is performed can be found elsewhere.[472]

Semi–parametric Cox proportional hazards regression

Cox proportional hazards regression is probably the most widely used form of survival regression. The method is called "semi–parametric" because, while it makes the proportional hazards assumption (see equations 3.7 and 3.8), with

$$\gamma = \exp(\mathbf{X}\beta) \tag{3.10}$$

the baseline survival and hazard functions are computed directly from the data. Furthermore, the survival curves depend only on the ranks of the survival times not the times themselves (though the method remains sensitive to skewed covariates). The estimated regression coefficients from a Cox model are the logarithms of the hazard ratios associated with those variables, so by taking the exponential of the regression coefficients, estimates of the hazard ratios are obtained. Cox proportional hazards regression is often preferred to fully parametric methods of survival analysis (such as exponential or Weibull models) as it gives very efficient (hazard ratio) estimates compared with parametric proportional hazards models. The key assumption of the Cox model, namely the proportional hazards assumption, may be assessed through examination of a smoothed plot of the Schoenfeld residuals against time.[475]

Time dependent covariates

In certain circumstances, it may become necessary to fit "time–dependent covariates": (1) when the covariates themselves change over time (e.g. an individual may receive a coronary angioplasty during the period of follow–up which would greatly influence their subsequent risk of CHD); and (2) when the regression coefficients vary over time (for instance the influence of a risk factor may become more or less important as one gets older). When either, or both, of these are true, the proportional hazards assumption of the semi–parametric Cox model is violated and account needs to be taken of this "non–proportionality" in analyses. One way of doing this is to adopt the Cox regression approach using discrete time intervals for each individual. This approach is referred to as the Anderson–Gill approach[476] (which views the "time to first event" as a Poisson process with a low rate). Each row corresponds to a different follow–up interval, with independence between rows for the same individual being guaranteed through the "lack of memory" property of the Poisson process.

3.4.3 Floating absolute risks

Risk estimates from epidemiological studies are usually presented in terms of relative risk with respect to a reference exposure level. However, when the categorical risk factor has more than two levels, the confidence intervals for the non–reference categories contain a common component of variance due to random variation in the reference category. This can make it difficult to interpret whether or not there are any real differences in relative risk between any two "non–reference" levels. One solution to this problem is to present the relative risks as "floating absolute risks".[477] This does not alter their values but ascribes an appropriate variance to each group (rather than having variances for each group relative to one group that is arbitrarily chosen to have a relative risk of 1 and no associated error). This technique is particularly useful for graphical presentations, because it allows confidence intervals for the relative risk to be presented for *all* levels of the risk exposure, not just the non–reference levels. However, the use of floating absolute risks has proved controversial because of problems of interpretation.[478–482] In this thesis, floating absolute risks are therefore calculated for the purpose of displaying relative risks estimates in figures only, with tables displaying the "normal" confidence intervals for each level relative to the single reference category. The method used to calculate the floating absolute risks can be

found in Easton and Peto.[477]

3.4.4 Re–sampling techniques

The jack–knife technique

When both estimation of risk associations and prediction of absolute risk are required from a single sample of data, it is usual to partition the data into two groups: a "training" group that is used to estimate regression parameters (to build the model); and a "prediction" group that is used to predict risks. The reason for this is that if the same sample were used in both steps (i.e. if the linear predictor were applied to the same sample from which it was derived) then it is well known that the predicted risks would be biased, leading to over–estimation of true risk differentials in the population.[483;484] However, partitioning the sample into a training group and a prediction group can lead to a loss of information both in the estimation of the regression coefficients and prediction of event risks, and may even lead to other sources of bias being introduced. An alternative method of estimation and prediction that uses all the data is provided by the "jack–knife" technique.[485] This method predicts the risk of failure for the ith individual from regression coefficients $\boldsymbol{\beta}_{[-i]}$ derived from the regression model that includes all subjects except the ith. Though this method of prediction is computationally intensive (N different regression procedures, each of which are based on $N-1$ individuals, need to be fitted), the "jack–knife" predicted risks $\{p_{[1]},\ldots,p_{[n]}\}$ are both unbiased and efficient, and provide a valuable method of both estimating associations and predicting disease risks from a single sample.

Bootstrap resampling and bias–corrected percentiles

Bootstrap resampling is a method for statistical inference often used to estimate standard errors and confidence intervals for statistics when the sampling distribution of that statistic is not known. For instance, suppose that independent and identically distributed observations X_i, $i=1,\ldots,n$ are observed, and that one wishes to estimate a parameter that can be defined as some function $\theta = T(x)$ of the values in the population, which is estimated from the data by the statistic $\hat{\theta} = T(X)$. Providing that the observed data are representative of the underlying data, bootstrap resampling may be used to empirically approximate the unknown cumulative distribution function of θ. This is done by repeatedly sampling (with replacement) from the data to obtain a large number (m) of bootstrap

replicate data sets (at least 1000 replications is usually recommended), and then evaluating the function for each replicate to give a set of bootstrap values $\{\hat{\theta}_i^B\}$, $i = 1, \ldots, m$. The empirical distribution of these bootstrap values $\hat{F}_b(X)$ approximates the theoretical sampling distribution $F_\theta(x)$. Therefore, approximate $100(1-\alpha)\%$ confidence intervals for $\hat{\theta}$ may be obtained by taking the $\alpha/2$ and $1-\alpha/2$ quantiles of $\hat{F}_b(X)$ as the lower and upper limits.[473] When the sample size is not very large however, the empirical percentiles of $\hat{F}_b(X)$ may not be very accurate. One method of adjusting for this is to calculate *bias corrected percentiles*, which take into account any differences between the estimate of the statistic from the original data ($\hat{\theta}$) and the median of the estimated values from the bootstrap sample $\{\hat{\theta}_i^B\}$.[473] This provides an improved estimate of the percentiles over the raw (observed) bootstrap percentiles.

3.5 Within–person variability in continuous risk factors

3.5.1 Introduction

In order to describe the effects of within–person variation in continuous coronary risk factors on estimated disease relationships, the differences between an individual's "baseline" risk factor level and an individual's "true" or "usual" level over a period of time need to be defined. The "baseline" risk factor level is simply the level that is measured at a single point in time (usually at the start of a prospective study). In contrast, the true or usual measurement level is the "long–term" or average risk exposure level for that individual over a certain period of time. For each person in a prospective study, this period will be defined as the interval over which the individual is "at risk" of CHD. In prospective studies of cardiovascular disease, it is the relationship between these usual or long–term risk factor levels and disease risk that is truly of interest. In practice, the accepted method of estimating these relationships is to use baseline levels as estimates of usual levels, and to perform analyses that examine the relationships between these levels and the risk of disease over the following period. However, baseline measurements of risk factors may differ from usual levels for a number of reasons:

1. Random measurement error, due to one or more of laboratory error, observer error or subject–recall error.

2. Short–term but true deviations from an individual's usual level. For example, an individual's usual heart rate in any one day would be unlikely to be estimated appropriately if measured immediately after vigorous exercise, as the rate would be temporarily high.

3. Long–term changes in an individual's risk factor level.

These collective influences are hereon referred to as sources of "within–person" variation. In order to describe the effects that within–person variation can have on the assessment of true relationships between risk factors and disease risk, it is helpful to introduce the following simple notation for the case of a single continuous risk factor subject to within–person variation. Consider a prospective study consisting of N individuals, whose observed baseline risk factor levels are denoted by $\{x_1, \ldots, x_N\}$. For each individual, the observed baseline level x_i may be assumed to be related to an unobserved usual value z_i through the equation

$$x_i = z_i + \varepsilon_i \tag{3.11}$$

where ε_i is random within–person variation having mean zero. Furthermore, let the mean exposure level in the population be $\mu = E[z_i]$, the true between–person variance be $\sigma^2 = \mathrm{Var}(z_i)$ and the within–person variance be $\tau^2 = \mathrm{Var}(\varepsilon_i)$, and assume that within–subject variation is independent of the true measurement level so that the variance of the observed measurement level is $\mathrm{Var}(x_i) = \sigma^2 + \tau^2$. Finally, denote the intraclass correlation coefficient (the ratio of the true between–person variance to the observed variance) by:

$$\lambda = \frac{\sigma^2}{\sigma^2 + \tau^2} \tag{3.12}$$

3.5.2 Regression dilution bias

In order to assess the influence of z_i on another normally distributed variable y_i, the usual approach would be to perform a least squares linear regression analysis with z_i as the predictor variable as follows:

$$y_i = \alpha^* + \beta^* z_i + \delta_i, \qquad \text{where } \delta_i \sim N(0, \phi^2) \tag{3.13}$$

However, as z_i is not actually observed, it is usually replace by its "baseline" estimate x_i. Providing that z_i, ε_i and δ_i are independently distributed it is known that the regression of y_i on x_i is linear as follows:

$$y_i = \alpha + \beta x_i + \gamma_i, \qquad \text{where } \gamma_i \sim N(0, \psi^2) \tag{3.14}$$

where

$$\beta = \beta^* \lambda \tag{3.15}$$

Therefore, by regressing y_i on the baseline levels x_i rather than the true average levels z_i, the estimated "baseline" regression coefficient β underestimates the true regression coefficient β^* by a factor equal to $1/\lambda$. This bias has long been recognised in linear regression,[486–488] and has also been shown to occur with logistic regression[16;489] and also Cox proportional hazards regression.[490;491] In the case of a single continuous risk factor, provided that the random within–person variation ε_i is unbiased and independent of the underlying usual value and of the response variable, the regression coefficient is always underestimated. This phenomenon is therefore often referred to as "regression dilution bias", while the correction factor λ is known as the "regression dilution ratio",[17] a term used throughout this thesis.

Regression dilution bias and the null hypothesis

Within–person variation results in underestimation of a risk factor's association with disease risk. However it is important to note that, under the assumptions stated in section 3.5.1, correction for within–person variation should not affect the statistical significance of a risk factor's association. The reason for this is that the null hypotheses $H_0 : \beta^* = 0$ and $H_0 : \beta = 0$ are equivalent and so they should have the same rejection

regions. Confidence intervals for adjusted regression coefficients should therefore preserve the significance (or lack of significance) of the observed baseline association.

3.5.3 Estimating the regression dilution ratio

In order to adjust for the effects of regression dilution bias, information regarding the regression dilution ratio λ is required. This can most easily be estimated from repeated measurements of risk factors, and a variety of methods are available for its estimation.[20] A brief review of some of the main methods is now provided. For each of these methods it is assumed that x_{ij} represents the jth repeat measurement for the ith individual. Furthermore, it is assumed that individuals with repeated measurements are either separate from the main study used to estimate β, or else form a small subset of the main study (so that the estimates of β and λ are essentially independent). If the data are from an independent source then it is important that they are also representative of the population under study. The descriptions that follow are based on the situation where two observed measurements are available for each individual (the "baseline" and "follow–up" measurements), which are denoted by $\{x_{11}, \ldots, x_{N1}\}$ and $\{x_{12}, \ldots, x_{N2}\}$ respectively. A complete review of the various methods available to estimate λ, as well as a discussion of their relative efficiency, is provided elsewhere.[20]

MacMahon's non–parametric method

This method was developed by MacMahon *et al.*[17] to correct for the effects of within–person variation in a baseline blood pressure measurement when estimating the subsequent relationship with the risk of stroke and coronary heart disease. The method divides the baseline sample into equal groups (usually fifths) and uses the fact that for each group the mean follow–up measurement $\bar{x}^k_{\cdot 2}$ $(k = 1, \ldots, 5)$ is an unbiased estimate of the true mean follow up level $\bar{z}^k$. The correction factor λ can therefore be estimated by comparing the range in means at follow–up with the range at baseline as follows:

$$\hat{\lambda}_{\mathrm{MAC}} = \frac{\bar{x}^5_{\cdot 1} - \bar{x}^1_{\cdot 1}}{\bar{x}^5_{\cdot 2} - \bar{x}^1_{\cdot 2}} \tag{3.16}$$

The method is appealing because it shows how the regression dilution bias may be thought of as a result of "regression to the mean", the phenomenon that reflects the fact that extreme values are, on average, more extreme than their true underlying values. However, MacMahon's method has several flaws. First, the choice of five groups is arbitrary: repeating the analysis with divisions based on quartiles or deciles would be equally valid but would lead to different estimates in practice. Second, by using the mean levels in the top and bottom categories only, the information provided by the middle three categories is lost. Though this could be overcome by using regression to estimate the slope of the relationship between the baseline group means and the follow–up group means, the logical extension of this (whereby individual data is used, rather than group means) forms the basis for the following method.

Rosner's regression method

MacMahon's method uses the fact that, conditional on x_{i1}, x_{i2} is an unbiased estimate of z_i, and so the relationship between x_{i2} and x_{i1} can be used to estimate λ. This principle also underlies the regression method proposed by Rosner *et al.*[16] Consider the special case of equation 3.13 where $y_i = x_{i2}$. It follows that $\beta^* = 1$ because the observed follow–up measurements x_{i2} deviate from the true levels z_i only by random within–person variation. Therefore the regression relationship between x_{i2} and x_{i1} is:

$$x_{i2} = \alpha + \lambda x_{i1} + \gamma_i, \qquad \text{where } \gamma_i \sim N(0, \psi^2)$$

and hence an unbiased estimate of λ can be obtained from the usual least squares regression slope

$$\hat{\lambda}_{\text{REG}} = \frac{n\sum x_{i1}x_{i2} - \sum x_{i1}\sum x_{i2}}{n\sum x_{i1}^2 - (\sum x_{i1})^2} = \frac{\sum(x_{i1} - \bar{x}_{\cdot 1})x_{i2}}{\sum(x_{i1} - \bar{x}_{\cdot 1})^2} \tag{3.17}$$

Confidence intervals for $\hat{\lambda}_{\text{REG}}$ can be calculated using the usual formulae for least squares regression.

Correlation coefficient methods

The "regression–based" methods considered so far depend only on the ordering of the baseline and follow–up measurements (and for Rosner's method also that the usual assumptions of linear regression are satisfied). However, if the baseline and follow–up samples have the same distributional assumptions then the decision to regress the second measurement on the first is arbitrary and an equally valid correction factor would be obtained by regressing the first on the second. Since the correlation coefficient is the geometric mean of these two coefficients, the correction factor could also be estimated as:

$$\hat{\lambda}_r = \frac{\sum(x_{i1} - \bar{x}_{i1})(x_{i2} - \bar{x}_{i2})}{\{\sum(x_{i1} - \bar{x}_{i1})^2 \sum(x_{i2} - \bar{x}_{i2})^2\}^{1/2}} \tag{3.18}$$

This is because the (product moment) correlation coefficient above has very similar properties to the intraclass correlation coefficient (in fact, it is the same when each pair of observations are counted twice, the second time in reverse).[492]

3.5.4 Correction for regression dilution bias

Equation 3.15 illustrates that in order to correct for regression dilution bias in a single continuous risk exposure, one simply needs to divide the observed association β by the estimate of the regression dilution ratio. An equivalent method of estimating the true coefficient β^* is obtained by "shrinking the data towards the mean" by calculating the conditional expectation of z_i given x_i which, under the assumptions of section 3.5.1 is equal, to

$$w_i = E[z_i \mid x_i] = \hat{\mu} + \lambda(x_i - \hat{\mu}) \tag{3.19}$$

β^* may then be estimated as the coefficient obtained through the regression of y_i on w_i as follows:

$$y_i = \alpha' + \beta' w_i + \theta_i, \qquad \text{where } \theta_i \sim N(0, \eta^2) \tag{3.20}$$

It can easily be shown that this "regression calibrated" coefficient β' is exactly the same as β^* obtained through equation 3.15. One advantage of using this approach in practice is that there is no restriction on the form that the relationship between the observed baseline level x_i and the expected true usual level w_i may take. In particular, the functional relationship described in section 3.5.1 may be extended to allow for "higher–order relationships" (e.g a quadratic relationship) between the baseline level and the expected usual level. This technique was utilized in the analyses of the influence of blood pressure on vascular disease risk performed by the Prospective Studies Collaboration in 2002,[113] where the authors estimated that the relationship between baseline blood pressure x and the usual blood pressure 3.4 years later F was quadratic,

$$F = 145.9 + 0.669(x - 150) - 0.0017(x - 150)^2$$

3.6 Within–person variability in categorical risk factors

3.6.1 Introduction

Categorical risk factors (such as cigarette smoking) are equally susceptible to the effects of within–person variation as continuous risk factors. In the past, many researchers have addressed the effects of within–person variation for categorical risk factors in terms of estimating the extent and effects of "misclassification" of exposure.[22–24;493] However, the term "misclassification" is generally used to describe only the "random error" component of within–person variation, and not the other components described in section 3.5.1. In this thesis, the definition of misclassification is extended to include changes in risk factor level as well as random error in risk factor assessment. A review of the established methods for correcting for within–person variation ("misclassification") in categorical risk factors is now provided.

3.6.2 "Misclassification" of a single dichotomous risk factor

In the simplest case of a dichotomous (binary) risk factor, the likelihood of an individual being "misclassified" based on a baseline screening is usually defined in terms of "sensitivity" and "specificity":

$$\begin{aligned} \text{Sensitivity} &= Pr(\text{Subject observed to have the risk factor when they truly do}) \\ \text{Specificity} &= Pr(\text{Subject not observed to have the risk factor when they truly don't}) \end{aligned}$$

If the risk factor is recorded without error then both the sensitivity and the specificity are equal to 1. However when one or more of the sensitivity and specificity are less than one, observed associations between the risk factor and disease risk underestimate the true associations (as for single continuous risk factors). In the special case whereby the probability of a false positive (1 - specificity) is equal to the probability of a false negative (1 - sensitivity), it can be shown that the overall probability of being classified correctly, denoted by α, may be estimated from repeated data through the equation

$$\hat{\alpha} = \frac{1}{2} + \frac{1}{2}\left(\frac{N-2n}{N}\right)^{1/2} \tag{3.21}$$

where N individuals are measured twice and there are n disagreements between the first and second measurements.[23] This estimate of the likelihood that the information provided by a baseline assessment of a binary risk factor is "correct" allows associations between the risk factor and the occurrence of disease to be adjusted. Denoting the log odds of disease (equivalently the log hazard) obtained from the use of baseline data in analyses by $\hat{\beta}$, the true log odds ratio β^* may be estimated by

$$\hat{\beta}^* = \frac{\hat{\beta}}{2\hat{\alpha}-1} \tag{3.22}$$

3.6.3 Categorical factors with more than two levels

The method of correcting an association for misclassification of a binary risk factor is not easily generalisable to the case of categorical factors that have more than two exposure levels. In this section, two methods of taking into account within–person variation in categorical factors that have more than two levels are described (note that these methods are also equally appropriate for continuous factors). These approaches differ from the

methods so far described for dealing with within–person variation in a single continuous risk factor in that the effect of the adjustment on the estimated disease relationship cannot necessarily be predicted. The goal of these methods is to use information obtained from repeated risk factor measurements to derive exposure levels that are in some way "superior" at predicting disease risk than baseline exposures alone. For each of the methods described, it is assumed that repeated follow–up measurements of the risk factor are available at various points over the study period.

Fitting time updated effects

As described in section 3.4.2, repeated measurements of an individual's risk factors *during a study* may be taken into account in a survival analysis by fitting "time–updated covariates". An example of where it may be of particular interest to fit such effects would be in an assessment of the effects of cigarette smoking on long–term disease risk. A prospective study may compare the incident disease rates between those who were active smokers at baseline with those who who non– or ex–smokers at baseline, however these risk differences may not truly quantify the risks from *continual* smoking if a substantial number of active smokers quit during the study (as these individuals may have been more likely to develop the disease if they continued). However, if it were known at what point the individuals gave up smoking, then providing that this was *before they developed disease* (i.e. while they were still "at risk"), a time updated effect could be included which would automatically "switch" individuals from "current smokers" to "ex–smokers" at the appropriate time. In this way, the effects of *continual* smoking on disease risk throughout the study could truly be evaluated. This approach would not estimate the effect on disease risk of giving up smoking, but it would take into account the fact that this occurs when making estimates of the effects of continual current smoking. The method is generalisable to known changes in any risk factor (not just categorical factors), but does require knowledge of how much and, more importantly, when these changes occur.

Cumulative "average" exposures

If frequent repeated measurements of risk factors are available during the study period then, rather than using the baseline risk level in analyses, one could use the cumulative average exposure to that factor over the study period. For instance, if blood pressure

were recorded every week throughout a prospective study then, for each individual, the average blood pressure per week "at risk" could be evaluated and related to disease risk in preference to the baseline measure. This is likely to provide a blood pressure assessment that is substantially nearer the true long–term average blood pressure for that individual than their baseline level. This method of using "cumulative" exposures may be generalised to categorical exposures, particularly ordinal categorical variables. In chapter 5 of this thesis, this "averaging approach" to risk factor assessment is applied to physical activity level and alcohol intake (which are defined an ordinal variables: none, occasional, light, etc.) and is also applied to the assessment of the average number of cigarettes smoked "per day at risk of CHD" for individuals defined as current smokers at baseline.

3.7 Multivariate correction for within–person variation

3.7.1 Introduction

The methods described in sections 3.5 and 3.6 use information from repeated measurements of risk factors to help better determine relationships between "usual" risk factor levels during the study and disease risk over the same period. In the case of a single continuous risk factor linearly associated with disease risk, it was shown how the risk of disease is underestimated when single baseline measures are used in analyses. However, when one or more covariates are subject to within–person variation, it does not necessarily follow that multivariate relationships are underestimated to the same degree as would be estimated when considering them in isolation. Indeed, it has been shown that under such circumstances observed regression coefficients may either under or overestimate true regression coefficients,[26;27] particularly if the variables are highly correlated.[27] Furthermore, risk relationships for other factors, even those with no within–person variation may be inaccurately assessed. In this section, a review of the methods for correcting for within–person variation in multiple covariates is provided. A more detailed description of these methods may be found in Rosner *et al.*[26]

3.7.2 Rosner's multivariate correction method

Rosner's multivariate correction method generalises the method described in section 3.5.2 where the effects of within–person variation in a single continuous risk factor were con-

sidered. Comparable notation to that used in section 3.5.2 is used now to describe the multivariate method.

Let X be a $k_1 \times 1$ vector of observed normally distributed exposure variables subject to within–person variation and Z be the associated $k_1 \times 1$ vector of true observations. Furthermore, define U to be a $k_2 \times 1$ vector of exposure variables (continuous or categorical) that are not subject to within–person variation, and P to be probability of a particular dichotomous disease outcome D. We suppose that the probability of disease is related to the true risk exposure variables Z and U through the vectors of true regression coefficients $\beta^*_{(1)}$ and $\beta^*_{(2)}$ in the following (logistic) regression model:

$$\ln\left(\frac{P}{1-P}\right) = \alpha^* + \beta^*_{(1)}Z + \beta^*_{(2)}U$$

and that the probability of disease is related to the observed risk exposure variables X and U through the vectors of observed regression coefficients $\beta_{(1)}$ and $\beta_{(2)}$ via

$$\ln\left(\frac{P}{1-P}\right) = \alpha + \beta_{(1)}X + \beta_{(2)}U$$

Now define the vector of true regression coefficients by $\mathbf{B}^* = (\beta^*_{(1)}, \beta^*_{(2)})$ (where $\beta^*_{(1)}$ is the vector corresponding to the true log odds ratios of D on Z (after controlling for U) and $\beta^*_{(2)}$ is the vector corresponding to the true log odds ratios of D on U (after controlling for Z)), and suppose that in addition to having information on the observed exposure levels X and U, an external sample of repeated measurements of Z (denoted by X_1 and X_2 for the first and second measurements), as well as a corresponding sample of U, are available. Defining $\lambda_{(1)}$ and $\lambda_{(2)}$ to be the $k_1 \times k_1$ and $k_1 \times k_2$ matrices of regression coefficients obtained through the multivariate linear regression of X_2 on X_1 and U obtained from the model

$$X_2 = \alpha' + \lambda_{(1)}X_1 + \lambda_{(2)}U + e \tag{3.23}$$

where α' is a vector of intercept terms and e is an error vector vector which follows a

multivariate normal distribution, the true regression coefficients $\mathbf{B}^*$ may be estimated from the observed regression coefficients $\hat{\mathbf{B}} = (\hat{\beta}_{(1)}, \hat{\beta}_{(2)})$ through the equation

$$\hat{\mathbf{B}}^* = \hat{\mathbf{B}}\hat{\Lambda}^{-1} \tag{3.24}$$

where

$$\hat{\Lambda} = \begin{pmatrix} \hat{\lambda}_{(1)} & \hat{\lambda}_{(2)} \\ 0_{k_2 \times k_1} & I_{k_2 \times k_2} \end{pmatrix} \tag{3.25}$$

This method leads to approximately unbiased estimate of $\mathbf{B}$ providing that the conditional distribution of the true exposure Z on the observed exposures (X and U) is the same for the main and external study data and also that the distribution of the observed exposure X given the true exposure Z is the same for both diseased and non–diseased subjects.

3.8 Overview

This chapter provides an overview of the statistics commonly used to quantify a risk factor's importance, reviews some of the main regression methods employed in analyses of prospective studies, and provides an overview of the effects that within–person variation in risk factors can have on estimated disease relationships (both in a univariate and a multivariate setting) as well as the methods with which one may take account of these effects. In particular, it is noted that within–person variation in continuous risk factors tends to lead to underestimation of a risk factor's importance (regression dilution bias) when these estimates are derived from baseline risk factor levels. Similarly, for binary risk factors, misclassification of exposure can lead to underestimation of the true risks associated with the factor. Established methods to correct for these effects are available provided that repeated risk factor measurements are available. Methods to take into account within–person variation in categorical exposures with more than two levels are also available. When multiple risk factors subject to within–person variation are related simultaneously to disease risk however, estimated baseline associations may under or overestimate true disease relationships. Under such circumstances, extra care should be taken to explore the

nature and effect of the correlation structure between the variables.

Chapter 4

The British Regional Heart Study

4.1 Summary

The British Regional Heart Study (BRHS) is a prospective study of cardiovascular disease in one General Practice in each of 24 British towns. Between 1978 and 1980, 7,735 men aged 40–59 years were recruited into the study (78% response). These men were representative of all middle–aged British men in terms of social class and were representative of all major regions at that time. Since the baseline assessment, surviving men have attended for follow–up examinations after approximately 16 years (in two towns only) and after 20 years (in all towns). Postal follow–up questionnaires, sent to study participants in 1983–85, 1992, 1996 and 1998–2000, and completed by a high percentage of participants, have provided information on lifestyle risk factor changes and medication use over the entire follow–up period. Major coronary heart disease events including non–fatal myocardial infarction and deaths from coronary heart disease have been ascertained from a combination of National Health Service central registers and two yearly reviews of General Practice records. The British Regional Heart Study has obtained high participant response rates throughout its history and losses to follow–up have been minimal. In this chapter, a broad overview of the design and methods of the BRHS is provided; more specific details pertinent to the work in this thesis are described in the relevant results chapters. In addition, descriptive statistics for the number and causes of death observed after twenty years of follow–up are provided.

4.2 Background

4.2.1 Aims

The British Regional Heart Study[61] was established in 1978 to determine: (i) the occurrence, natural history and management of cardiovascular disease in British men; (ii) the individual determinants of cardiovascular disease; and (iii) the factors responsible for the considerable geographic variations in cardiovascular mortality in Great Britain. The first phase of the BRHS was an ecological study which examined cardiovascular and other mortality over a five year period around the 1971 census from 253 towns in England, Scotland and Wales.[494] The original study hypothesis was that factors operating at a town level such as water hardness, climate and air pollution may have been responsible for the variations in disease observed between different towns across Great Britain. Twenty four of these towns subsequently formed the basis of the prospective phase of the BRHS, the selection process for which is described below.

4.2.2 Selection of study towns

Study towns were selected according to the following general principles: (i) they represented all standard regions in Great Britain; (ii) each town had a population at the 1971 census between 50,000 and 100,000 and was separate from major conurbations; (iii) geographic variations in cardiovascular mortality and water hardness were represented; (iv) the towns were representative of the region in terms of socio–economic factors; (v) towns with recent large housing developments, noticeable population movement (migration) or unusual population structures were excluded; and (vi) the towns should include some of the towns that were 'outliers' when mortality from cardiovascular disease was plotted against water hardness (e.g. Hartlepool, Exeter and Harrogate). Towns that met these general criteria were randomly selected for inclusion into the prospective phase of the BRHS. In order to provide strong geographic representation, the size criteria was relaxed in specific cases – those of Ipswich, whose population was 122,700 at the 1971 census, and the Scottish towns of Dunfermline, Ayr and Falkirk, whose populations were below 50,000. The 24 towns selected for participation in the main study are shown in figure 4.1 and table 4.1.

4.2.3 Selection of general practices

In order to increase the likelihood of a good initial response and good follow–up rates, as well as to facilitate the organisation and administration of the field work, it was decided that study participants would be recruited from one group practice in each town. The criteria for choosing the practice included its size (over 7,500 patients and two or more general practitioners) and its representativeness of the socio–economic profile of the town population. Potential practices were then sent information about the study and asked about their interest and willingness to participate. After visiting each practice, one group practice from each town was invited to participate in the study. If no age and sex register existed for the practice, one was prepared by the study team. This was required for 19 of the 24 practices.

4.2.4 Selection of study participants

From the age and sex register of each practice, 450 men aged 40–59 stratified by 5–year age groups were selected at random. Individuals for whom it was felt (by the general practitioner) that they would be unable to participate due to severe mental or physical disability were excluded. It was emphasised that men with cardiovascular disease should not be excluded unless there was another reason to do so. Exclusions accounted for some 6 to 10 men per practice. The remaining subjects were invited to participate in the study in a letter signed by the practice doctors. Response rates were high, ranging from 70% to 85% between the different towns (see Table 4.1).

4.3 Baseline examination

The baseline assessment of study participants took place between January 1978 and June 1980. A team of three nurses visited each town in turn. The assessment consisted of: (i) an administered questionnaire; (ii) physical measurements; (iii) a resting electrocardiogram (ECG); and (iv) a blood sample. The nurses were trained in order to standardize procedures, including administering the questionnaire, before the study and at regular intervals throughout the period of baseline data collection. A summary of the main methods used at the baseline examination is shown in Table 4.2, a detailed description of which now follows.

4.3.1 Questionnaire

The baseline questionnaire (shown in Appendix B) was administered at the screening site by one of the nurses. The questionnaire covered demographic details, family history, socio–economic details, cardiovascular and respiratory symptoms, medical history and lifestyle risk factors. The areas of the questionnaire particularly relevant to the work in this thesis are summarised below:

1. **Chest pain:** The World Health Organisation (WHO) Rose chest pain questionnaire was used with some modifications to assess presence and intensity of current angina symptoms. The modifications included a reordering of the questions:

   (a) use of the present tense rather than the past tense (i.e. "Have you ever had any pain...", was replaced with, "Do you ever have any pain..."),

   (b) inclusion of two additional questions eliciting frequency of occurrence of chest pain, and

   (c) rephrasing of the two exertion questions (subjects were asked whether walking or hurrying 'produced' the chest pain rather than simply whether they got chest pain when walking or hurrying).

   Current chest pain was categorised as: (i) no chest pain; (ii) non–exertional chest pain, defined as chest pain not brought on by walking or hurrying; and (iii) angina, defined as chest pain brought on either by walking uphill or hurrying, or by walking at an ordinary pace on the level. The latter group was subdefined as *definite* angina if the pain caused the subject to stop or slow down in response to the pain, if the pain was relieved on stopping, if relief occurred within 10 minutes and if the site of the pain included the sternum or left anterior chest (sites 4, 5 or 8 in Figure 4.2), and as *possible* angina if at least one of these criteria was satisfied. In addition, definite and possible angina were classified as *grade I* if the chest pain was brought on only by walking uphill or hurrying, and as *grade II* if it was brought on by walking at an ordinary pace on the level.

2. **Severe chest pain:** Analysis of a separate question together with a 'site of pain' diagram allowed further assessment of possible myocardial infarctions. This was defined as severe chest pain lasting half an hour or more, situated in the sternum

or left anterior chest (sites 4, 5 or 8 in Figure 4.2), that caused the participant to consult a doctor.

3. **Recall of doctor diagnoses:** Participants were asked about whether they had ever been told by a doctor that they had had (among other things): angina, a heart attack (with alternatives of 'myocardial infarction' and 'coronary thrombosis'), a stroke or diabetes. Recall of a diagnosis of either angina or a heart attack was subsequently defined as recall of doctor diagnosed CHD.

4. **Current treatment:** The men were asked which treatments they were currently taking that were prescribed by a doctor. These included specific questions on the use of blood pressure lowering and lipid lowering drugs.

5. **Social class and geographic region:** Men were asked about their longest held occupation in terms of type, designation and status. On the basis of this, social class was determined using the Registrar General's classification (I, II, IIINM, IIIM, IV, V or Armed Forces; see Table 4.3).[495] Men not in the armed forces were subsequently classified as non–manual (I, II and IIINM) or manual (IIIM, IV and V). Geographic region was also defined according to study town, and grouped as: (i) Scotland; (ii) northern England; (iii) Midlands and Wales; and (iv) southern England.[496] These regions are shown in Figure 4.1.

6. **Lifestyle risk factors:** ***Cigarette smoking status***, defined as current smokers, ex–smokers or non smokers, was ascertained from an adapted form of a smoking questionnaire developed in 1970 by the Medical Research Council Social Medicine Unit, London School of Hygiene and Tropical Medicine. Those who smoked less than 1 cigarette per day were classified as non–smokers (or ex–smokers if they had previously regularly smoked more than 1 cigarette per day). ***Alcohol intake*** was ascertained from questions inquiring about frequency, quantity and type of alcoholic beverage consumed, and was categorised into eight groups: (1) lifetime tee–totallers; (2) one to two times a month or on special occasions; (3) weekend drinkers (one to two drinks per day); (4) weekend drinkers (three to six drinks per day); (5) weekend drinkers (more than six drinks per day); (6) daily drinkers (one to two drinks per day); (7) daily drinkers (three to six drinks per day); and (8) daily drinkers (more than six drinks per day). These categories were then re–classified into the following

five groups: (i) lifetime tee–totallers (1 above); (ii) occasional drinkers (2 above); (iii) light drinkers (3, 4 or 6 above); (iv) moderate drinkers (5 and 7 above); and (v) heavy drinkers (8 above). ***Physical activity*** during work and leisure time was assessed by a number of questions regarding regular walking or cycling, recreational activity and sporting activity. Based on the frequency and type of activity, a physical activity score was derived for each man,[193] from which a six–level index was created: (i) inactive; (ii) occasional (regular walking or recreational activity); (iii) light (more frequent recreational activities or vigorous less than once a week); (iv) moderate (cycling, very frequent recreational activities or sporting activity once a week); (v) moderately vigorous (sporting activity at least once a week or frequent cycling, plus frequent recreational activities or walking, or frequent sporting activity only); and (vi) vigorous (very frequent sporting exercise or frequent sporting exercise plus other recreational activities). Baseline lifestyle characteristics by study town are shown in Table 4.4.

4.3.2 Physical examination

Height was measured without shoes using a Harpenden Stadiometer with digital meter which recorded to the nearest millimetre. Weight in trousers and socks was measured to the nearest 0.1 kg using an MPS110 filed survey scale (beam balance). Body mass index (BMI) was subsequently calculated as weight in kilograms divided by height in metres squared (kg/m^2). Blood pressure was measured twice in succession in the right arm, with the subject seated and the arm supported, using the London School of Hygiene and Tropical Medicine sphygmomanometer. Diastolic blood pressure was recorded at the disappearance of Korotkoff sounds (phase V). Blood pressure readings were subsequently adjusted for observer variation within each town[497] and the mean of the two blood pressure measurements on each individual was used in analyses. Lung function was measured using a Vitalograph spirometer; forced expiratory volume in one second (FEV1) was subsequently standardised for height. Baseline physical measurements by study town are shown in Table 4.4.

4.3.3 Electrocardiogram (ECG)

A resting electrocardiogram was recorded on computer tape using the three–lead orthogonal system[498] and viewed on an oscilloscope for any major abnormalities. The completed tapes for each town were analysed by computer in the Department of Medical Cardiology in Glasgow. On the occasions where technical problems occurred, 12–lead electrocardiograms were recorded on a Hewlett–Packard (1515–B) machine and sent for analysis. The validity of using the three–lead system and the interpretation of results have been previously reported.[499;500] Ischaemic abnormalities were classified into one of four groups. *Definite MI* was defined as a broad Q wave in any lead, together with a Q–R amplitude ratio greater than 1:3 if the Q wave was in the anterolateral (X) or inferior (Y) leads. This equates closely to Minnesota codes 1–1 and 1–2–1 to 1–2–6.[499] *Possible MI* was defined as above but was dependent on the width of the Q wave and the magnitude of the Q–R amplitude ratio. This equates with Minnesota codes 1–2–7, 1–2–8 and 1–3. *Definite myocardial ischaemia* and *possible myocardial ischaemia* required a combination of ST segment depression and T wave changes (negativity or low positivity). The degree of ST depression required for abnormality is less than that required by the Minnesota code, although flat or downward sloping ST segment is additionally required in the three–lead system. Definite ischaemia equates with Minnesota code 4–1 or 5–1, and possible ischaemia with 4–1, 4–2, 5–2, 5–3.

4.3.4 Blood measurements

Blood samples were taken in the non–fasting state between 0800 and 1830 hours, separated and transferred overnight to central laboratories. Estimation of serum total cholesterol and HDL cholesterol was carried out at the Wolfson Research Laboratories, Birmingham. Serum total cholesterol was measured by a modified Liebermann–Burchard method on a Technicon SMA 12/60 analyser and HDL cholesterol by the Liebermann–Burchard or enzymic procedures after precipitation with magnesium phosphotungstate.[501] For 18 towns only (excluding Harrogate, Shrewsbury, Lowestoft, Mansfield, Southport and Merthyr Tydfil), triglycerides were measured using an enzymic method at the Department of Chemical Pathology, Royal Free Hospital, London. For men in these towns, LDL–cholesterol was subsequently calculated using the Friedrickson–Friedwald equation: LDL cholesterol = Total cholesterol - HDL cholesterol - (0.45 × Triglycerides). Measurements of base-

line lipids were subsequently adjusted for the effects of time of day.[502] Serum insulin was measured at the Department of Diabetes and Metabolism, University of Newcastle, using an enzyme–linked immunosorbent assay (ELISA) which does not cross–react with proinsulin.[503] Glucose was estimated by a Technicon SMA 12/60 analyser at the Wolfson Research Laboratories, Birmingham. For men in 18 towns, baseline levels of cotinine were analysed in 2001 using a gas–liquid chromatography method (detection limit 0.1 ng/ml).[504] Baseline biochemical measurements by study town are shown in Table 4.4.

Nested case–control studies of new risk factors

Between 1998 and 2000 several nested case–control studies were carried out on the BRHS data in order to examine the potential influence on CHD risk of a range of new risk factors including homocysteine, C–reactive protein and various haemostatic factors including von Willebrand factor, fibrin D–dimer and t–Pa antigen.[356;358;505;506] Cases were selected from those men that had had a major CHD event before 1996, and were frequency matched with controls based on age and town. For analysis of C–reactive protein, von Willebrand factor, fibrin D–dimer and t–Pa antigen, approximately 600 cases and 1,200 controls were selected,[356;358;505] for homocysteine 386 cases and 454 controls were selected.[506] Baseline blood samples were then thawed and reanalysed for these new risk factors. Details of the methods and laboratories used to estimate these factors are given in section 4.4.4.

4.3.5 Evidence of pre–existing CHD

Several criteria were used to assess whether study participants had experienced any coronary heart disease prior to entry into the study. These included:

1. recall of doctor diagnosis of angina or myocardial infarction based on the questionnaire.

2. World Health Organization (Rose) chest pain questionnaire evidence of definite or possible angina (grade I or grade II). [i]

[i]This relaxed criteria has a high sensitivity for identifying men with angina (when compared with the "gold standard" of expert clinical opinion) while maintaining a reasonable degree of specificity.[507–509] Though some men will inevitably be identified as "false positives", since the purpose of this assessment is to identify all those men with any evidence of CHD (with a view to excluding them from analyses) an inevitable trade–off in terms of specificity was deemed acceptable.

3. a history of severe chest pain lasting half an hour or more that caused them to consult with a doctor (possible myocardial infarction).

4. ECG evidence of definite or possible myocardial infarction or ischaemia.

Table 4.5 shows the responses of the study participants to each of these assessments and how these answers varied by geographic region. 15% of all study participants had symptoms of or recalled a diagnosis of myocardial infarction or angina (using criteria 1 – 3), and a further 10% had ECG evidence of definite or possible myocardial infarction or ischaemia, so that one quarter of all men had some evidence of coronary heart disease. This figure was lower in the South of England than elsewhere. Few individuals (< 1%) recalled a doctor diagnosis of stroke.

4.4 Follow–up procedures

Since the baseline assessment of study participants, all men have been systematically followed for fatal and non–fatal cardiovascular events for over twenty years. Information on mortality has been ascertained from National Service registers in Southport (England and Wales) and Edinburgh (Scotland) and information on cardiovascular morbidity has been ascertained from a notification system supplemented by two–yearly reviews of general practice records. At regular periods since the baseline examination in 1978–80, surviving men have also been asked to complete questionnaires enquiring about their health and lifestyle characteristics. These questionnaires were completed by study participants after an average of 5, 13, 17 and 20 years of follow–up for each individual, as shown in Figure 4.3. In 1996, a study investigating the interrelationships between cardiovascular risk factors, clinical disease, intima media thickness and carotid plaque was carried out in Dewsbury and Maidstone (two towns with widely differing CHD rates), during which physical measurements and blood samples were remeasured.[510] After 20 years of follow–up for all individuals (between 1998 and 2000), all surviving men were invited to attend for rescreening, at which physical and biochemical measurements were measured. In order to estimate short–term within–subject variability in both established and novel risk factors, physical and biochemical measurements were repeated one week apart on an age–matched sample from a local general practice in Islington, North London. The measurements taken for this study were exactly the same as for the 20–year rescreening of the BRHS partici-

pants. A detailed review of the follow–up procedures, questionnaires and examinations is now provided.

4.4.1 Mortality follow–up

Information on death was collected through the established "tagging" procedures provided by the National Health Service Central Registers (NHSCR) in Southport (England and Wales) and Edinburgh (Scotland). Copies of death certificates were sent to the study centre at three–monthly intervals, and included identification details, place, date and cause of death, the name of the certifier, and whether a post mortem examination was performed. A fatal CHD event was defined as a death with ischaemic heart disease (International Classification of Diseases (ICD) 9th revision codes 410–414) as the underlying cause including sudden death of presumed cardiac origin. In cases of contradictory or inconsistent causes of death, clarification was sought with the general practitioner, hospital consultant, or pathologist.

Of the 7,735 men enrolled into the study, 2,077 (26.9%) died of all causes within the first 20 years. Table 4.6 shows a breakdown of these deaths. Ischaemic heart disease was the single most common cause of death over 20 years, accounting for 36% of all deaths occurring, marginally higher than all cancer related deaths combined (35%). Stroke contributed a further 6% of observed deaths, and other circulatory disease a 7%, so that approximately half of all deaths were due to diseases of the circulatory system. For only 13 deaths was the cause unknown, accounting for 0.6% of all deaths.

4.4.2 Cardiovascular morbidity follow–up

Record review process

On entry to the study, the medical record of each man was stamped and a blue card inserted which could be completed by the general practitioner and returned to the study centre in the event of a new diagnosis of one of: (i) myocardial infarction; (ii) stroke; (iii) angina; or (iv) transient ischaemic attack. Additionally, each practice co–ordinator carried out two yearly reviews of study participants' medical records, using a listing supplied by the study centre. The practice co–ordinator updated the list with new deaths, emigrations and removals, and reviewed each subject's medical records in order to identify all new cardiovascular events and diagnoses occurring in the preceding two years. In the event of

an uncertain diagnosis, copies of medical records and/or hospital letters were forwarded to the study centre for interpretation. After the first eight years of follow–up (between 1986 and 1988) a complete retrospective review of study participants' notes over the preceding eight years was carried out in order to check diagnoses to that date. The blue card system was subsequently withdrawn but the two yearly record review has continued ever since. Men who registered with a new general practitioner in Britain subsequent to the baseline examination were tracked through the registration procedures at the Family Health Service Authorities and, where necessary, through the NHSCR. After twenty years of follow–up, fewer than 1% of study participants had been lost to follow–up.[511]

Assessment of cardiovascular morbidity events

Evidence regarding non–fatal heart attacks were obtained from the two yearly reviews of the patients' notes, including hospital and clinic correspondence, as described above. A non–fatal heart attack was diagnosed according to established World Health Organization criteria, which included any report of myocardial infarction accompanied by at least two of the following: a history of severe chest pain, electrocardiographic evidence of myocardial infarction, and cardiac enzyme changes associated with myocardial infarction. Throughout this thesis "major" coronary heart disease events are defined as non–fatal myocardial infarction or death from coronary heart disease.

4.4.3 Follow–up questionnaires

Figure 4.3 shows the dates that study participants completed the various health and lifestyle questionnaires. Following the initial baseline questionnaire, three additional questionnaires were sent to all surviving men who were still resident in Britain. The first was sent at the fifth anniversary of each man's baseline assessment (Q5), the second in November 1992 (Q92) and the third in November 1996 (Q96). A fourth questionnaire was completed by all surviving men who attended for re–examination after 20 years of follow–up (Q20). Men in Dewsbury and Maidstone who took part in the carotid–plaque substudy in 1996 completed their Q96 questionnaire in February and March of that year rather than in November. With the exception of Q20, all follow–up questionnaires were self administered. A maximum of two reminders (with additional questionnaires and reply–paid envelope) were sent to non–responders. Response rates for the follow–up postal question-

naires were 98% for Q5, 90% for Q92 and 88% for Q96. Of men who survived 20 years, 77% attended for examination and completed the Q20 questionnaire. Similar questions as were asked in the baseline questionnaire (including questions on lifestyle characteristics such as cigarette smoking and alcohol consumption) were also asked in each of the follow–up questionnaires. Level of physical activity was asked in all questionnaires except Q5, though additional questions about housing tenure and car ownership (which were not asked in the baseline questionnaire) were included in the Q5 questionnaire. In the twenty year follow–up questionnaire (Q20), participants were asked to record their current or most recent occupation (rather than their longest held occupation) and the duration of that employment. From this information occupational social class after 20 years was again derived using the Registrar General's classification.[442]

4.4.4 Follow–up screenings

Follow–up examination of study participants took place after approximately 16 years (for men in Dewsbury and Maidstone) and after 20 years (in all towns), as shown in Figure 4.3. The short–term variability study among men in Islington took place in 2000. Three observers made all measurements at the 20–year screening; one observer made all measurements at 16–year screening and the short–term variability study. At all screenings, men were asked to attend for examination between 0800 and 1800; for the short–term variability study this was at the same time on both occasions. Men were not asked to fast at the 16–year screening; at the 20–year screening and the short–term variability study, all men not taking insulin or oral hypoglycaemic treatment for diabetes were asked to fast for a minimum of six hours, during which they were instructed to drink only water. At all examinations, physical measurements including height, weight and blood pressure were recorded, and blood samples (taken using the Sarstedt Monovette system) were collected for measurement of a range of established and novel coronary risk factors. These samples were separated and frozen at –20°C on the day of collection and transferred in batches to central laboratories. A review of the methods used to ascertain these measures at each screening is now provided.

- **Physical measurements:** Height and weight were recorded to the nearest millimetre and 0.1 kg respectively. Blood pressure was measured twice in succession in the right arm, with the subject seated and the arm supported, using a Dinamap 1846

oscillometric blood pressure recorder; over–reading of systolic pressure by the instrument[512] was corrected in analysis. At the 20–year screening, blood pressure readings were adjusted for observer variation within each town.[497] The mean of the two blood pressure measurements on each individual was used in analyses throughout.

- **Blood lipids, insulin and glucose:** Blood lipids and glucose were analysed at the Department of Chemical Pathology, Royal Free Hospital, London (Prof. A Winder, Dr M Thomas); insulin was measured at the Department of Diabetes and Metabolism, University of Newcastle (Prof. KGMM Alberti, Ms P Shearing). Serum total cholesterol was measured using the method of Siedel[513] and HDL cholesterol using the method of Sugiuchi,[514] both on a Hitachi 747 automated analyser which also analysed triglyceride concentrations. LDL–cholesterol values were again calculated using the Friedrickson–Friedwald equation. Serum insulin was measured using the same method used at baseline; plasma glucose was measured using the method of Trinder[515] using a Falcor 600 automated analyser. At the 20–year screening, LDL cholesterol, triglycerides, insulin and glucose levels were adjusted for the effects of time of day and time since last meal.[516]

- **Haemostatic and inflammatory variables and homocysteine:** Haemostatic and inflammatory variables were measured in citrated blood plasma at the Department of Medicine, University of Glasgow (Prof. GDO Lowe, Dr A Rumley). Fibrinogen was measured using the Clauss method. Plasma levels of t–PA antigen and D–dimer were measured with enzyme linked immunosorbent assays as was von Willebrand factor (vWF) antigen. C–reactive protein was assayed by ultra sensitive nephelometry. Serum total homocysteine was determined using a modified automated assay, based on pre–column derivatisation with monobromobimane, followed by reverse phase high performance liquid chromatography with fluorescence detection, and was measured at the Department of Pharmacology, University of Bergen (Prof. H Refsum, Prof. P Ueland).

Table 4.1: Towns included in the British Regional Heart Study

Town	SMR for CVD in men aged 35–64	Population size (1971)	Number of men examined	Response rate (%)
Ayr	140	47,890	301	70
Bedford	80	72,880	298	73
Burnley	114	76,130	287	80
Carlisle	121	71,820	389	85
Darlington	109	85,900	382	82
Dewsbury	142	51,130	325	79
Dunfermline	118	48,890	352	80
Exeter	90	93,800	332	84
Falkirk	98	37,600	309	75
Gloucester	84	89,980	311	73
Grimsby	96	95,610	318	71
Guildford	78	58,090	335	82
Harrogate	82	63,470	280	77
Hartlepool	101	97,110	313	77
Ipswich	92	122,700	362	85
Lowestoft	85	52,120	324	83
Maidstone	99	71,250	318	72
Mansfield	95	57,820	321	80
Merthyr Tydfil	135	55,100	283	76
Newc–Under–Lyme	115	77,320	293	77
Scunthorpe	109	70,900	332	76
Shrewsbury	95	56,630	311	83
Southport	114	84,870	322	80
Wigan	134	81,140	337	77

SMR = standardised mortality ratio.

Table 4.2: Summary of key measurements taken at the baseline examination

Measurement	Method
Current chest pain	WHO (Rose) chest pain questionnaire. Categorised as: (i) no chest pain; (ii) non–exertional chest pain; and (iii) angina. Angina subdefined as definite or possible depending on severity of symtoms, and as grade I or grade II depending on the conditions under which the symptoms occur.
Severe chest pain	Rose angina questionnaire together with 'site of pain' diagram (Figure 4.2). Defined as symptoms lasting half an hour or more requiring consultation with a doctor.
Recall of diagnosis	Ever told by a doctor that they had had (among other things): angina, a myocardial infarction, a stroke or diabetes
Electrocardiogram	Three lead system. Ischaemic abnormalities classified according to Minnesota criteria into four groups, definite and possible MI, and definite and possible myocardial ischaemia.
Blood presssure	Average of two seated measurements using London School of Hygiene and Tropical Medicine sphygmomanometer. Diastolic blood pressure recorded at the disappearance of Korotkoff sounds (phase V).
Serum total cholesterol	Modified Liebermann–Burchard method on a Technicon SMA 12/60 analyser.[501]
HDL cholesterol	Liebermann–Burchard or enzymic procedures after precipitation with magnesium phosphotungstate.[501]
Triglycerides	Enzymic method (18 towns only).[501]
LDL cholesterol	Estimated from total, HDL and triglycerides through the Friedrickson Friedwald equation (18 towns only).
Serum insulin	ELISA assay that does not react with proinsulin.[503]
Glucose	Measured using a Technicon SMA 12/60 analyser.
Cigarette smoking	Several questions used to ascertain cigarette smoking exposure, including type and amount of tobacco smoked, and the number of years since started (or quit) smoking.
Physical activity	Several questions used to gauge type of activity and frequency. Categorised as: (i) inactive; (ii) occasional; (iii) light; (iv) moderate; (v) moderately vigorous; and (vi) vigorous.[193]
Alcohol intake	Several questions used to gauge frequency, quantity and type of alcoholic beverage. Categorised as: (i) lifetime tee–totallers; (ii) occasional drinkers; (iii) light drinkers; (iv) moderate drinkers; and (v) heavy drinkers.[265]

Table 4.3: Registrar General's six category classification of social class

Class	Description	Examples
I	Professional	Doctors, lawyers, scientists
II	Managerial and technical occupations	Managers, teachers, white–collar workers
IIINM	Skilled occupations (non–manual)	Nurses, shop assistants
IIIM	Skilled occupations (manual)	Electricians, plumbers
IV	Partly skilled occupations	Bus drivers
V	Unskilled	General labourers, cleaners
	Armed forces	Army, Navy, Air Force

Table 4.4: Selected mean baseline characteristics by study town of the 7,735 men in the British Regional Heart Study

Town	Mean TC	Mean HDL	Mean SBP	Mean DBP	Mean BMI	Mean Height	Current smokers (%)	Physically active (%)	Heavy drinker (%)	Manual (%)
Ayr	6.27	1.22	143.4	81.2	25.1	171	50.8	35.1	11.3	64.5
Bedford	6.08	1.11	148.0	85.0	25.4	174	27.9	37.9	5.0	47.9
Burnley	6.40	1.14	146.0	84.5	25.1	172	45.8	22.3	14.0	69.9
Carlisle	6.58	1.19	149.9	88.2	25.3	173	40.6	31.7	11.1	59.5
Darlington	6.45	1.14	146.6	84.0	25.3	174	33.8	42.0	11.0	43.0
Dewsbury	6.44	1.20	150.8	82.7	25.5	173	50.5	28.9	13.8	65.2
Dunfermline	6.30	1.09	152.4	88.5	25.4	173	45.9	42.8	7.1	63.7
Exeter	6.51	1.13	138.9	78.5	25.8	174	37.9	40.4	4.8	47.5
Falkirk	6.21	1.13	147.9	85.5	26.1	172	49.2	37.1	11.3	74.8
Gloucester	6.05	1.15	144.8	81.2	25.9	172	44.8	38.6	9.0	72.6
Grimsby	6.24	1.15	148.4	85.9	25.6	172	60.4	27.8	19.2	84.0
Guildford	6.27	1.17	135.9	77.6	24.8	176	24.2	45.5	3.6	23.8
Harrogate	6.41	1.22	138.6	82.6	25.8	175	31.8	44.2	10.0	35.7
Hartlepool	6.11	1.14	147.7	85.8	25.6	173	42.5	29.9	26.8	74.6
Ipswich	6.28	1.15	142.6	79.3	25.5	174	32.0	46.7	5.0	42.8
Lowestoft	6.50	1.12	142.2	76.0	25.2	174	37.2	43.3	3.1	64.7
Maidstone	6.26	1.12	146.3	83.4	25.5	174	43.1	33.8	8.8	56.5
Mansfield	6.43	1.15	143.7	79.0	25.7	174	40.6	37.0	8.4	60.1
Merthyr Tydfil	6.19	1.18	148.8	82.1	25.6	171	47.7	27.2	15.9	72.5
N–U–L	6.27	1.12	149.0	82.1	25.7	173	48.1	34.4	14.0	68.4
Scunthorpe	5.99	1.07	140.4	78.2	25.8	173	48.8	41.8	10.5	78.9
Shrewsbury	6.45	1.19	135.8	77.4	25.3	174	33.9	45.1	8.0	41.7
Southport	6.22	1.12	147.2	82.0	25.3	174	36.4	44.5	7.5	44.3
Wigan	6.15	1.17	147.9	82.4	25.4	173	39.9	30.4	21.1	65.5
All towns	**6.30**	**1.15**	**145.2**	**82.2**	**25.5**	**173**	**41.3**	**37.2**	**10.8**	**59.1**

Table 4.5: Percentage of men with baseline evidence of cardiovascular disease by study area

Criterion	South $n = 2280$	Midl/Wales $n = 1208$	North $n = 3285$	Scotland $n = 962$	Total $n = 7735$
Recall of doctor diagnosis of myocardial infarction or angina	3.5	7.1	5.8	7.0	5.5
Rose angina questionnaire evidence of definite or possible angina	5.4	9.1	8.5	9.9	7.9
History of severe chest pain	7.9	10.2	9.4	9.6	9.1
Any of above	**11.9**	**16.6**	**15.9**	**17.4**	**15.0**
ECG evidence of definite or possible myocardial infarction or ischaemia	13.6	16.1	14.7	14.6	14.6
Any of above	**22.2**	**26.9**	**26.2**	**26.3**	**25.1**

Table 4.6: Causes of death over 20 years of follow–up in the BRHS

Cause of death		n	(%)
1.	Ischaemic heart disease	751	36.2%
2.	Malignant neoplasms	729	35.1%
3.	Diseases of the respiratory system	151	7.3%
4.	Other diseases of the circulatory system	142	6.8%
5.	Stroke	129	6.2%
6.	Diseases of the digestive system	46	2.2%
7.	Disease of the nervous system and sense organs	22	1.1%
8.	Endocrine, nutritional and metabolic diseases	15	0.7%
9.	Diseases of the genitourinary system	14	0.7%
10.	Diseases of blood and blood–forming organs	9	0.4%
11.	Mental disorders	6	0.3%
12.	Diseases of the musculoskeletal system and connective tissue	4	0.2%
13.	Other causes	46	2.2%
14.	Unknown cause	13	0.6%
Total number of deaths		2,077	100%

Figure 4.1: Twenty–four towns of the BRHS split into four areas

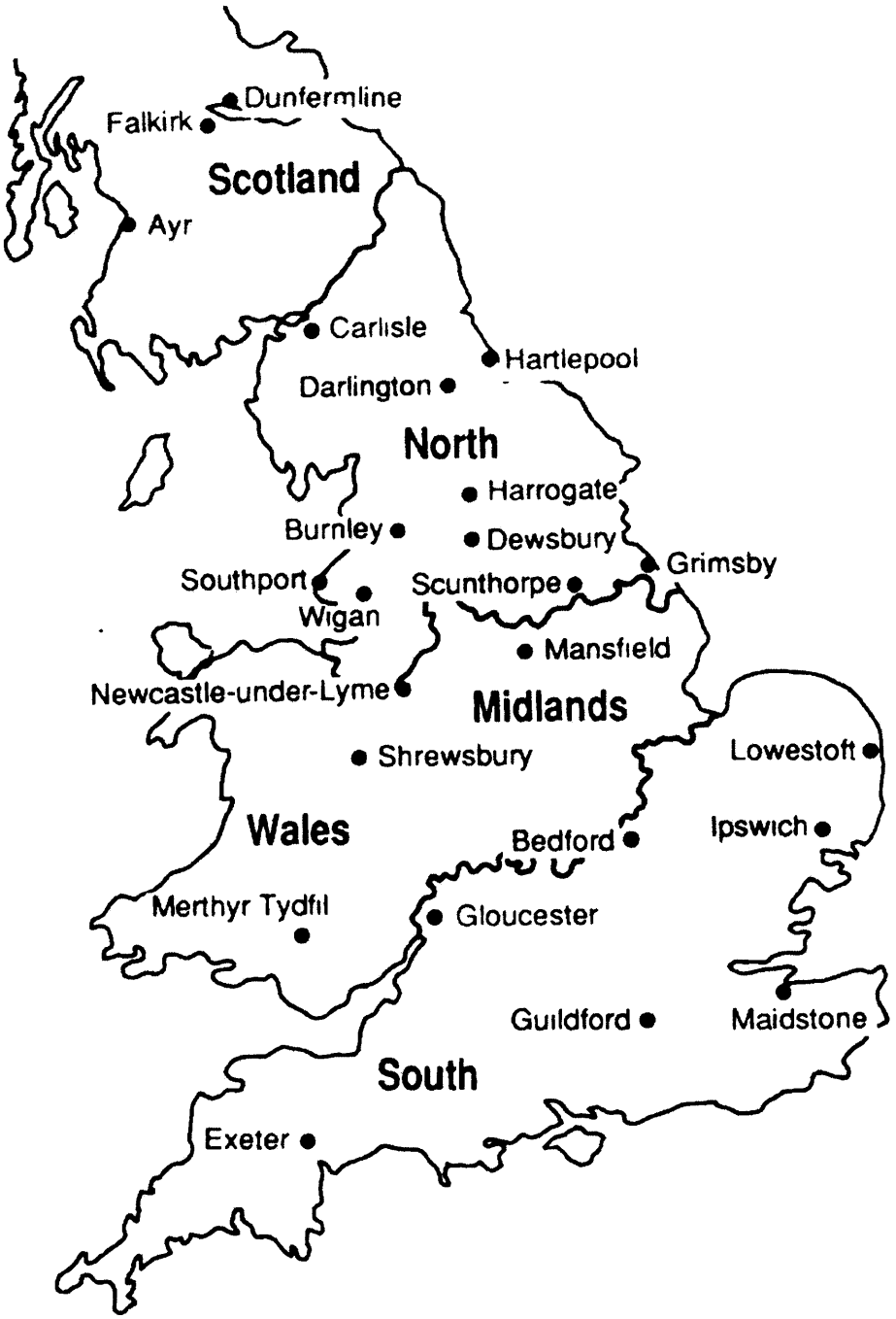


Figure 4.2: Chest pain chart shown to study participants at baseline

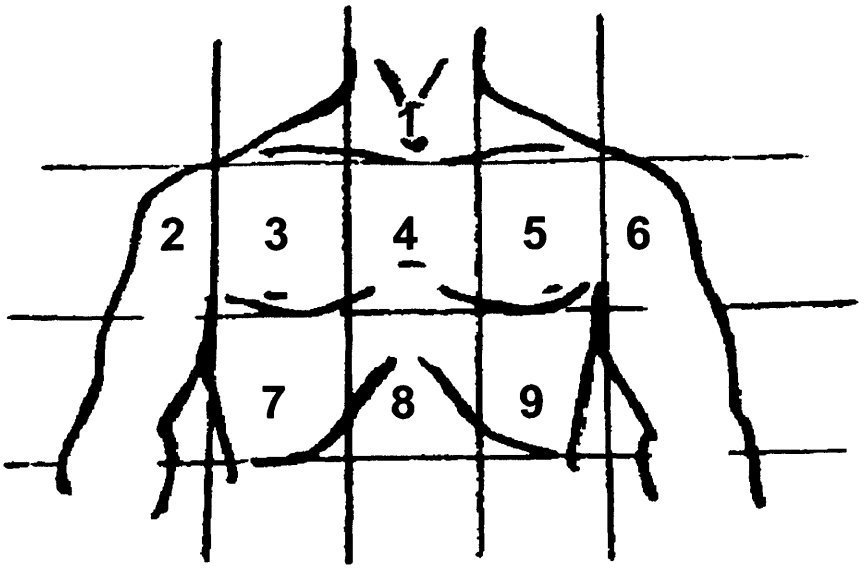
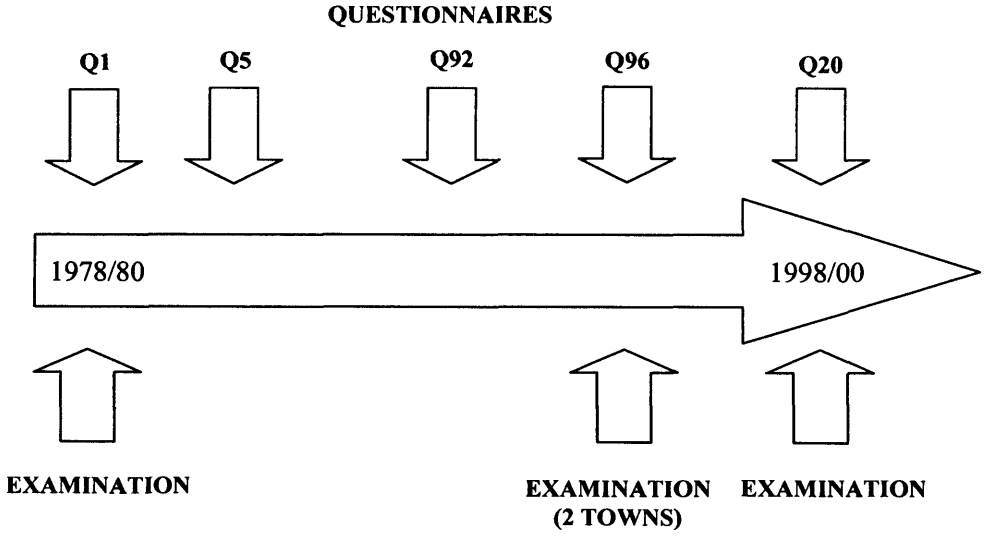


Figure 4.3: Follow–up in the British Regional Heart Study



Chapter 5

Extent of within–person variation in CHD risk factors

5.1 Summary

In this chapter the repeated measurements available in the British Regional Heart Study are used to examine the extent of within–person variation for established and novel CHD risk factors. For the physical and biochemical factors, this is examined by estimating regression dilution ratios over periods of one week, four years, sixteen years and twenty years. The influences of age, disease severity, treatment and social class on the regression dilution ratios are examined, and the effect that regression dilution in one factor may have on risk associations for other risk markers, including both precisely and imprecisely measured markers, is assessed. For the categorical risk factors, including cigarette smoking and physical activity, the extent of within–person variation is estimated using information provided by the follow–up questionnaires. The results show substantial levels of within–person variation in both continuous and categorical risk factors over the study period. For continuous risk factors, the use of baseline measures of both established and novel risk factors in analyses results in marked underestimation of their associations with disease risk, the extent of which increases with duration of follow–up. For example, the use of baseline measurements to estimate associations with CHD risk 20 years later results in underestimation by an estimated 47% (95% confidence interval (CI) 44% to 50%) for serum total cholesterol, 51% (95% CI 48% to 55%) for systolic blood pressure, and 76% (95% CI 73% to 78%) for diastolic blood pressure. In a multivariate setting where the

independent relations of a blood lipid and a blood pressure measure (and other factors) are to be assessed, the use of total cholesterol and systolic pressure was found to result in the least residual bias in estimation of other risk associations. For the categorical risk factors, a substantial amount of within–person variation was observed over the study period. In particular, there was a tendency for baseline levels to overestimate the proportion of the population at either end of the risk exposure distribution, particularly the proportion of the population who are truly heavy smokers or heavy drinkers.

5.2 Introduction

5.2.1 Background

In chapter 3 it was described how within–person variation in coronary risk factors (random error, short term biological variability and long term variability) can lead to inaccurate assessments of risk relationships. For single continuous risk factors, it was shown that these relationships are underestimated when derived from observed baseline measurements (a phenomenon referred to as regression dilution bias). This bias exists because the differences between subjects estimated from a single baseline measurement exaggerate the true differences that exist between those subjects over a period of time. Therefore any variations in risk that are attributed to the risk exposure at baseline should really be attributed to a narrower range of values. The effects of regression dilution bias are particularly relevant to coronary heart disease, as several of the key established and novel risk factors are subject to within–person variation over time. Previous reports have tended to examine these effects as they apply to measures of blood cholesterol and blood pressure, two of the key determinants of CHD risk.[17;19;517–522] Long–term follow–up of individuals in the Whitehall and Framingham studies, for instance, has led to the conclusion that true associations between total cholesterol and blood pressure with CHD risk are underestimated by around one third during the first decade of follow–up, one half during the second and by as much as two thirds during the third decade of follow up.[19]

However, with the exception of these studies, there is relatively little published information on the extent of regression dilution bias at different follow–up intervals for established, and particularly for novel, coronary risk factors. Furthermore, the extent of within–person variation in categorical risk factors has been little studied and the validity

of performing univariate correction methods in a multiple covariate setting is uncertain.

5.2.2 Objectives

The aim of this chapter is to assess the extent of within–person variation in both established and novel CHD risk factors in middle–aged British men followed for 20 years, and the likely effect that this would have on estimated disease relationships in both a univariate and a multivariate setting. For the continuous coronary risk factors, the regression dilution ratio over a variety of follow–up periods (one week, and four, sixteen and twenty years) is estimated for the established risk factors (blood lipids and blood pressure), a range of novel potential risk factors (including glucose, insulin, fibrinogen, fibrin D–dimer, homocysteine, C–reactive protein and Von Willebrand factor, all of which have been associated with coronary heart disease risk),[300;343;356;358;505;523] as well as other potentially relevant haemostatic measurements (factor VII and tissue plasminogen activator (t–PA)). A secondary objective is to assess whether the effects of regression dilution bias depend on other factors (including age, town, presence of disease and social class), in order to assess the potential generalisability of these effects to other populations. For the categorical risk factors, the extent of within–person variation is estimated using information from the four follow–up questionnaires (Q5, Q92, Q96 and Q20) to derive "average" exposure levels to the factors, in order to allow comparisons with the baseline classifications.

5.3 Methods

5.3.1 Data sources

The extent of within–person variation in physical and biochemical risk factors was assessed by estimating regression dilution ratios over various interval periods using the following data:

1. Baseline data (1978–80); all physical measurements, blood lipids, insulin and glucose.
2. Sixteen–year Dewsbury Maidstone screening (1996); all physical and biochemical measurements.
3. Twenty–year screening (1998–2000); all physical and biochemical measurements.

4. Nested case–control studies of new risk factors (1998–2000); subset of stored baseline samples from 18 towns were thawed and reanalysed for assessment of all new risk factors (except fibrinogen and factor VII).

5. One week repeatability study (Islington; 2000); all physical and biochemical measurements.

These data allow estimation of the regression dilution ratio for all physical and biochemical risk factors over periods of one week (Islington study), four years (Dewsbury/Maidstone 1996–2000), sixteen years (Dewsbury/Maidstone 1980–1996 plus nested case–control subgroup), and twenty years (all towns 1978/80 to 1998/00 plus nested case–control subgroup). For the categorical risk factors, information on each individual was taken from the baseline questionnaire together with (potentially) each of the four follow–up questionnaires (Q5, Q92, Q96 and Q20), in order to assess the extent of within–person variation in cigarette smoking, physical activity and alcohol intake over the study period.

5.3.2 Within–person variation in continuous risk factors

Estimating and interpreting the regression dilution ratio (RDR)

In the case of a single continuous risk factor subject to within–person variation, the RDR was estimated using Rosner's regression method described in section 3.5.3. This method estimates the RDR by the regression coefficient obtained from regressing the follow–up measurement on the baseline measurement, with 95% confidence intervals being calculated in the usual way for a regression coefficient. For variables that are normally distributed, the amount by which the "baseline association" underestimates the true association may be obtained by subtracting the RDR from 1, and multiplying by 100% (for instance, if the RDR calculated over a 5–year period was 0.8, then the association (say, the log odds ratio) between baseline levels and disease risk would underestimate the association between 5–year levels and disease risk by 20%). For variables where the underlying distribution is positively skewed, such as triglycerides, analyses were performed on the log scale. The interpretation of the RDR for these variables is the same, though applied to a variable on the log scale (it is the amount through which the association between the logarithm of the usual exposure and disease risk is underestimated because of the use of the logarithm of baseline levels in analyses).

Dependence of the RDR on the interval period

For each risk factor (except fibrinogen and factor VII where repeated data is only available over two periods), the relationship between the RDR and the interval over which it was calculated was approximated by fitting an inverse variance weighted negative exponential curve through the plot of the RDR against the interval period (see Figures 5.2 to 5.5). This was done by first regressing the logarithm of the RDR on the interval period (weighted by the inverse variances of the individual estimates) so that the "best" straight line relationship through the plot of log RDR against the interval could be obtained, and second, by transforming this straight line relationship back to the exponential scale (therefore, if the relationship between log(RDR) and the interval in years was expressed as, $y = a - bx$, the relationship between the RDR and the interval would be $y = \exp(a - bx)$).

Tests for heterogeneity

In order to test whether the regression dilution ratio corresponding to a particular risk factor over a certain interval period differed significantly between two or more groups of individuals (e.g. whether it differed by age group), the inverse variance weighted average of the separate estimates was first calculated,

$$L = \frac{\sum(\hat{\lambda}_i/s_i^2)}{\sum(1/s_i^2)} \tag{5.1}$$

where s_i is the estimated standard error of the ith estimate $\hat{\lambda}_i$, and then the sum of the squared (weighted) deviations from this statistic was calculated as shown below,

$$X^2 = \sum((\hat{\lambda}_i - L)/s_i)^2 \tag{5.2}$$

This statistic was then compared with a chi square distribution with $k - 1$ degrees of freedom in order to assess significance, where k is the number of subgroups being compared.

5.3.3 Within–person variation in categorical risk factors

Using the baseline data together with data from (potentially) each of the four follow–up questionnaires, "average" exposures to cigarette smoking, alcohol, and physical activity over the study period (or until the time to first CHD event if observed) were calculated for each individual. For cigarette smoking, men were defined as being either never smokers or ex–smokers *throughout the study* if they were defined as such at every questionnaire they responded to. For men who were current smokers at baseline, the number of cigarettes they reported to smoke at each questionnaire (zero if they had given up) was used to calculate the average number of cigarettes they smoked per day on the study. It was assumed that any changes in the number of cigarettes smoked between consecutive questionnaires occurred linearly over the intervening period (unless a CHD event occurred during that period in which case it was assumed that the individual continued to smoke at the same rate as reported at the last questionnaire prior to the date of that event). Men who were non–smokers at baseline (never or ex–smokers) but subsequently reported that they were current smokers were classified as "new/recurrent" smokers. For alcohol intake, a five–point scale was used to denote the intake level at the baseline assessment and each of the follow–up assessments (from 0 (none) to 4 (heavy); a six–point scale was used for physical activity). Using these data, average alcohol intake during the study (and average physical activity level) was calculated as indicated in Figure 5.1 (changes between questionnaires were again assumed to occur linearly unless the individual had a CHD event during that time). For instance, in the third scenario shown in Figure 5.1 where the subject has a non–fatal MI after 18 years of follow–up and has their risk exposure coded as '4' at baseline, '3' at Q92, '3' at Q96 and '2' at Q20 (and no response to Q5), the average exposure over the 18 years they were 'at risk of CHD' would be calculated as (42 (area of A) + 12 (area of B) + 6 (area of C))/18 = 3.3 (note that the information from Q20 is ignored because a CHD event was obserevd between Q96 and Q20 which may have subsequently caused in change in lifestyle). From these average exposure levels calculated for all 7,735 men, each subject was reclassified on the original categorical scale according to the nearest whole number (for example for alcohol intake, an average exposure of <0.5 was defined as "never drinker", 0.5 – 1.5 was defined as "light drinking" etc.).

5.3.4 Effects of within–person variation in more than one risk factor

In order to estimate the effect that within–person variation in more than one risk factor could have on estimated disease relationships in a multivariate setting, the inter–relationships between different blood lipid and blood pressure indices, and their relationships with other important coronary risk factors (age and cigarette smoking) were examined using the 20–year repeated data. In particular, in section 5.4.6 of this chapter, estimates of $\mathbf{\Lambda}$ (a matrix of regression coefficients obtained by regressing follow–up levels on baseline levels; see Chapter 3, section 3.7) are calculated to display the relationships between a blood lipid measure (total or HDL cholesterol), a blood pressure measurement (systolic or diastolic), age and cigarette smoking status (defined as never, ex– or current). However, because it is the inverse of $\mathbf{\Lambda}$ (and not $\mathbf{\Lambda}$ itself) that determines the extent to which baseline associations are modified (through the equation $\boldsymbol{\beta}^* = \boldsymbol{\beta}\mathbf{\Lambda}^{-1}$; see section 3.7), it is the inverse of $\mathbf{\Lambda}$ that is presented (this is referred to as the "modifying matrix"). The extent to which the off–diagonal components of $\mathbf{\Lambda}^{-1}$ deviate from zero indicate the extent to which univariate correction methods will be misleading. For consistency with later chapters, this particular analysis is restricted to men with no baseline evidence of CHD (symptoms, recall or doctor diagnosis of CHD).

5.4 Results

5.4.1 Characteristics of repeated data

Of the 7,735 men examined at baseline in 1978–80, 5,658 survived the following 20 years and 4,252 of these attended the 20–year re–screening. In Dewsbury and Maidstone, of 643 men examined at baseline, 532 survived to 1996 and 425 of these (80%) participated in the 16–year re–screening. Of the men who attended the Dewsbury–Maidstone 16–year screening in 1996, 400 survived to 2000 and 297 of these (74%) participated in the 20–year re–screening. The short–term variability study included 112 men who completed all measurements at two screenings taken a week apart. Data on regression dilution ratios for the established physical and biochemical risk factors after one week and four, sixteen and twenty years of follow–up (as well as estimates for the new risk factors over one week and four years) are therefore based on 112 men, 297 men, 425 men and 4,252 men respectively. In addition, the case–control studies performed between 1998 and 2000 that reanalysed

stored baseline samples provided further information on new risk factors over a 16–year period for 65 men (Dewsbury/Maidstone case–control participants that survived and were re–examined in 1996), and over a 20–year period for between 700 and 850 men (all case–control participants that survived and were re–examined in 1998–2000); though fewer (330 repeat samples) were available for homocysteine.

Table 5.1 shows the mean values of selected characteristics of the men who attended each of the screening examinations. For blood pressure, it is noticeable that for the short–term variability study the mean blood pressure (both systolic and diastolic) was lower at the second measurement than the first. Over 20 years, mean blood pressure increased by approximately 6 mmHg (systolic) and 4 mmHg (diastolic). Blood lipids were reasonably constant over short periods, but generally fell over 20 years (mean total cholesterol decreased from 6.3 mmol/L in 1980 to 6.0 mmol/L in 2000). Men who died before the 20–year examination or survived but failed to attend the 20–year examination had higher baseline systolic and diastolic blood pressure than surviving men that attended the 20–year examination (both $p < 0.0001$). However, baseline levels of triglyceride, total cholesterol, LDL cholesterol and HDL cholesterol did not differ significantly between these groups. For men in Dewsbury and Maidstone, there were no differences in baseline blood lipids (triglycerides and total, LDL and HDL cholesterol) or blood pressure levels between those men that attended both the 16 and 20–year screenings and those who did not.

5.4.2 Estimates of the regression dilution ratio for continuous factors

Tables 5.2 and 5.3 show estimates of the regression dilution ratio with 95% confidence limits for each of the measurements after 1 week and 4, 16 and 20 years of follow–up; these estimates are also displayed graphically in figures 5.2 to 5.5 where weighted regression has been used to identify the 'best' exponential fit to the estimates (for the new risk factors where only two estimates are available, the curve has been extrapolated to 20 years for illustrative purposes).

RDR estimates for blood lipids

For blood lipids, the effects of measurement error and short–term biological variation are apparent even over a one–week period; use of baseline measures as estimates of usual levels would result in short–term associations being underestimated by approximately 10–15%

for total cholesterol, LDL cholesterol and triglycerides. Associations with HDL cholesterol and the ratio of total to HDL cholesterol are subject to the least within–person variability, being hardly affected over this period. After four or more years, associations with usual levels would be underestimated by less than 15% for HDL cholesterol, by between 25–35% for total cholesterol, triglyceride and total:HDL cholesterol, and by around 40% for LDL cholesterol. For long–term associations, say those occurring 20 or more years after baseline, the effects of regression dilution bias when quantifying blood lipid associations are marked – use of baseline measures would result in true associations being underestimated by approximately one half for total cholesterol, LDL cholesterol and triglyceride. Associations with HDL cholesterol and the ratio of total to HDL cholesterol would be more accurately assessed, but would still be underestimated by, respectively, one quarter and one third.

RDR estimates for blood pressure

For blood pressure, the effects of regression dilution bias are large because of the greater degree of within–person variability associated with this measure. Even over a one week period, true associations would be underestimated by approximately 30% if single measurements were used in analyses. The magnitude of this underestimation increases sharply (particularly for diastolic pressure) as the interval period increases so that after four or more years, associations with usual blood pressure levels at that time would be underestimated by around 40% for systolic pressure and 50% for diastolic pressure. For mid blood pressure (the average of systolic and diastolic pressure) and mean arterial pressure (two thirds diastolic pressure plus one third systolic), the effects are between those for systolic and diastolic pressure. For long–term associations the effects of regression dilution for blood pressure are large; associations with systolic pressure would be underestimated by approximately one half (similar to that for total cholesterol), but for diastolic pressure, baseline associations would underestimate true associations by around three quarters.

RDR estimates for body mass index, insulin and glucose

Of all the risk factors presented, body mass index displayed by far the least within–person variation over the study period. Correspondingly, even over a 20–year interval, the effect of using baseline measures of body mass index as estimates of usual levels around that time would be virtually nil, possibly resulting in underestimation by around 7%.

For insulin and glucose, the effects of regression dilution bias were fairly modest over a one week period, but became considerably more marked from four years onwards. After four years, associations with usual insulin and glucose levels around that time would be underestimated by 50% or more. Over a 20–year interval period, these estimates increased to 73% for insulin and 67% for glucose.

RDR estimates for novel coronary risk factors

For the range of novel risk factors measured in the BRHS, including homocysteine, C–reactive protein and the haemostatic factors, repeat measurements were available over periods of one week, and four, sixteen and twenty years using a combination of data from the main study cohort as well as from the nested case–control studies performed between 1998 and 2000. For fibrinogen and factor VII only one week and four year data are available. Table 5.3 and Figures 5.4 and 5.5 show estimates of the regression dilution ratio for these factors over the various follow–up periods. Similar effects of regression dilution bias were observed as for the established risk factors; the regression dilution ratios for the novel risk factors are similar in magnitude to those of established risk factors, generally lying between those of total cholesterol and diastolic pressure. Over an interval period of at least four years, the use of baseline measures to estimate associations with usual exposure levels would result in underestimation by between 40–50% for fibrinogen and C–reactive protein, approximately 30% for Von Willebrand factor and t–PA, and approximately 25% for factor VII, D–dimer and homocysteine.

5.4.3 Factors influencing the regression dilution ratio

In order to assess the extent that estimates of the regression dilution ratio may be generalisable to other study populations, the 20–year estimates for the blood lipid and blood pressure indices were calculated separately according to age, baseline evidence of CHD (recall of doctor diagnosis of MI or angina, Rose questionnaire evidence of angina, or history of severe chest pain), social class (manual versus non–manual) and geographic location (the South, Midlands and Wales, the North, Scotland). These estimates together with χ^2 tests for difference are shown in Figures 5.6 to 5.11.

Age

In order to investigate whether the size of the regression dilution ratio varies with age we divided the study participants into four age groups at baseline: 40–44 years, 45–49 years, 50–54 years and 55–59 years. Separate RDR estimates for the primary blood lipid measurements (total, HDL, and LDL cholesterol, and the ratio of total to HDL) and blood pressure measurements (systolic and diastolic pressure) were then calculated for each of these groups over the following 20 years. For HDL and LDL cholesterol, as well as the ratio of total to HDL cholesterol and both blood pressure indices, no evidence of heterogeneity between the different estimates by age group was observed (all $p > 0.05$). For total cholesterol however, some evidence of heterogeneity was observed; the RDR for total cholesterol among men originally aged 55 to 59 years was slightly higher than that observed amongst other men, indicating that this group experienced slightly lower levels of within–person variation over the 20–year follow–up period.

Evidence of CHD and treatment

Of the 4,252 men for whom 20–year interval data were available, 452 (10.6%) had evidence of CHD at the baseline examination. The RDR estimates for the blood lipid and blood pressure measurements obtained from these men were generally slightly lower than the RDR estimates obtained from all remaining men (indicating potentially greater degrees of long–term within–person variation in these measurements experienced by individuals already with CHD), though these differences did not reach statistical significance. However, when individuals who developed CHD over the interval period were also included in the 'diseased' group, much larger differences in the RDR estimates were observed which did reach statistical significance. As would be expected, the 20–year RDR estimates for blood pressure differed markedly between the 1,139 men who had (at some time) received blood pressure lowering drugs and the 3,113 men who had not (0.37 versus 0.60 for systolic, 0.18 versus 0.30 for diastolic; both $p < 0.0001$). Furthermore, the 20–year RDR estimate for total cholesterol differed markedly between the 324 men receiving lipid lowering drugs at follow–up (predominantly statins) compared with all other men (0.35 compared with 0.64; $p < 0.001$). However, these differential effects did not explain the differences in the RDR estimates for blood pressure and blood cholesterol observed between those who did and those who did not develop CHD (interaction terms between the development of CHD and

the RDR remained highly significant $p < 0.002$ even after adjustment for drug use).

Social class and geographic location

Finally, estimates of regression dilution ratios over 20 years were compared by social class (manual versus non–manual) and geographic location (South England, Midlands and Wales, North England, and Scotland). Relatively little difference in the size of the regression dilution ratio by social class was observed, though some heterogeneity was observed for diastolic blood pressure, where the level of within–person variation over 20–years was marginally greater for manual men (possibly due to inter–relationships between social class and the development of CHD and subsequent treatment). For systolic and diastolic pressure, no geographic differences in the regression dilution ratio were observed. For total and LDL cholesterol, significant heterogeneity was observed between the different geographic regions; within–person variation tended to be greater in the North of England than the South, again possibly reflecting the higher incidence of CHD (and effects thereof) experienced in the North of England, though these differences could also simply reflect the role of chance.

5.4.4 Lifestyle risk factors and within–person variation

Table 5.4 shows the responses for cigarette smoking status (never, ex, current (1–20 a day), current (21–39 a day) or current ($\geq$40 a day)), physical activity (none, occasional, light, moderate, moderately vigorous or vigorous), and alcohol intake (none, occasional, light, moderate or heavy) at the baseline questionnaire (Q0) and each of the follow–up questionnaires (Q5, Q92, Q96 and Q20). It can be seen that cigarette smoking patterns recorded at each questionnaire changed considerably over the study period, due in part to real changes in smoking habits in the men and in part to the increased 'selectiveness' of the group as follow–up increases (due to the excess number of deaths amongst active smokers). These selection effects are reflected in the gradual increases that are observed in the proportion of men who are defined as 'never smokers' at each of the questionnaires, which increases from 24% at baseline to 29% at Q20. Of men who survived 20 years, only approximately 1 in 200 reported to be actively smoking at least 40 cigarettes a day, compared with 1 in 25 at the baseline examination. For physical activity, the modal exposure category at each of the follow–up questionnaires was 'occasional', comprising

nearly one third of the group at baseline, and approximately one quarter of the surviving cohort members at Q92, Q96 and Q20. The proportion of men reported to be 'inactive' at each questionnaire was fairly stable at around 10%; though a slightly higher proportion (15%) was reported at Q96, this could be artificially high because of the role of chance, rather than reflecting a true increase in physical inactivity in 1996 compared with other years. The proportion of men exercising 'vigorously' increased from 7% at baseline to 15% at Q20. For alcohol consumption, approximately one third of the men were categorised as 'light' drinkers at baseline, with a further one quarter of men categorised as 'occasional' drinkers and one quarter categorised as 'moderate' drinkers. During the follow–up period, there was a steady downwards trend in the amount of alcohol consumption reported by surviving study participants (due to a combination of the cohort becoming older and drinking less, secular changes in alcohol consumption with time, and possible survival 'selection' effects). After 20 years of follow–up, only 3% of the surviving men were classified as heavy drinkers (compared with nearly 11% at baseline), while 10% were classified as non–drinkers (compared with 6% at baseline).

5.4.5 Misclassification of "average" lifestyle risk exposures

For each individual, 'average' cigarette smoking group, physical activity level and alcohol intake during the individual's personal exposure period (either 20 years or the time of censoring/first CHD event, whichever is lower) was calculated as described in section 5.3.3, and compared with the baseline exposure. These comparisons are shown in Tables 5.5 to 5.7.

Misclassification of cigarette smoking

At baseline, 1,819 men were defined as never smokers (see Table 5.5). Of these, men 1,802 (99%) were confirmed as being never smokers when follow–up questionnaires were taken into account. The 17 men who were, at some later point, recorded as being current smokers were probably either due to random measurement error (by the study investigators) or subject recall error (misreporting by subject at baseline), though they may also reflect men who truly first began regular smoking during middle–age. Of the 2,715 men who were classified as ex–smokers at baseline, only a small proportion (7%) were subsequently categorised as active smokers. For the men who were active smokers at baseline, the

degree of 'misclassification' when compared with average smoking levels during the study varied considerably by the amount of tobacco reportedly smoked each day. Of those who initially reported smoking 1–20 cigarettes a day, the vast majority (96%) were classified as such when follow–up questionnaires were taken into account and classifications were based on average smoking exposure. For heavy smokers however, the baseline assessment alone tended to greatly overestimate the proportion of men that consistently smoked heavily during the study. Of men who (at baseline) reported smoking 21–39 cigarettes a day, less than half were classified as such after follow–up questionnaires were taken into account (the vast majority of the rest were classified as smoking 1–20 a day), while for the 316 men who initially reported to smoke at least 40 cigarettes a day, only 83 (26%) truly smoked to that degree *throughout the study.* Overall, the degree of misclassification in cigarette smoking status was modest: 6,739 men (87%) were classified the same based on baseline and average exposure levels.

Misclassification of physical activity

Reported physical activity levels at baseline and average levels throughout the study are shown in Table 5.6. At any given level of physical activity reported at baseline, only approximately 50–60% of the group were defined in the same category when follow–up questionnaires were taken into account. Most individuals (around 90%) were either defined the same based on average levels or else were within one category of their baseline exposure group. Overall, 4,531 men (57%) were categorised in the same group, but there was a tendency for baseline levels to overestimate both the proportion of the population that were inactive as well as the proportion that were highly active.

Misclassification of alcohol intake

Alcohol intake at baseline and average alcohol intake during the study are shown in Table 5.7. Within–person variation can be seen to increase with increasing alcohol intake, as reflected by the observation that 80% of those originally defined as non–drinkers were defined the same when based on average levels, compared with 65% of those originally defined as light drinkers and only 28% of those originally defined as heavy drinkers. Baseline assessment of alcohol exposure resulted in substantial overestimation of the number of men who were truly heavy drinkers during the follow–up period and underestimation

of the number of men who, on average, could be classified as non–drinkers.

5.4.6 Inter–relationships between risk measurements

To examine the effects that within–person variation in blood lipid and blood pressure measurements have on their estimated associations, as well as their effects on associations for age and cigarette smoking status (two factors that, for this analysis, are assumed to be measured precisely), the "modifying matrix" of regression coefficients $\mathbf{\Lambda}^{-1}$ over a 20–year period was calculated for the following four hypothetical scenarios:

1. an analysis that adjusts for total cholesterol, systolic blood pressure, age and cigarette smoking status.

2. an analysis that adjusts for total cholesterol, diastolic blood pressure, age and cigarette smoking status.

3. an analysis that adjusts for (log) HDL cholesterol, systolic blood pressure, age and cigarette smoking status.

4. an analysis that adjusts for (log) HDL cholesterol, diastolic blood pressure, age and cigarette smoking status.

In each scenario, estimates were restricted to men with no baseline evidence of CHD, and for each estimate of $\mathbf{\Lambda}^{-1}$, the rows/columns correspond to the variables presented in the following order: (1) blood lipid measurement; (2) blood pressure measurement (per 20 mmHg systolic or 10 mmHg diastolic); (3) age (years); (4) ex–smoker (1=yes, 0=no); and (5) current smoker (1=yes, 0=no). These four matrices are shown below (recall that the vector of true regression coefficients $\boldsymbol{\beta}^*$ is obtained from the vector of baseline associations $\boldsymbol{\beta}$ through the equation $\boldsymbol{\beta}^* = \boldsymbol{\beta}\mathbf{\Lambda}^{-1}$):

Estimates of $\mathbf{\Lambda}^{-1}$

$$\mathbf{\Lambda}_1^{-1} = \begin{pmatrix} 1.84 & 0.30 & 0.02 & 0.02 & 0.08 \\ 0.04 & 2.13 & -0.05 & 0.06 & 0.07 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{\Lambda}_2^{-1} = \begin{pmatrix} 1.88 & 0.64 & 0.04 & 0.08 & 0.18 \\ 0.46 & 4.19 & 0.06 & 0.37 & 0.51 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{\Lambda}_3^{-1} = \begin{pmatrix} 1.32 & 0.04 & 0.00 & 0.00 & -0.01 \\ -1.63 & 2.10 & -0.05 & 0.06 & 0.08 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{\Lambda}_4^{-1} = \begin{pmatrix} 1.31 & 0.05 & 0.00 & 0.00 & 0.00 \\ -1.56 & 3.99 & 0.05 & 0.38 & 0.47 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The matrices above show the extent by which the vector of "baseline" associations should be modified in order to obtain the vector of true associations between risk and risk factors 20–years after baseline. The diagonal components describe the extent to which each risk association is underestimated, because of the use of its baseline level in analyses, while the off–diagonal terms show the degree that these new associations are altered because of their relations with the blood lipid and blood pressure measure. For age and cigarette smoking (variables 3, 4 and 5 in the model), exposures are assumed to be known precisely and hence the diagonal components for these variables are equal to 1 (so that the baseline association is the true association). However, because of their relations with the blood lipid and blood pressure indices, their true regression coefficients are also modified by the estimated associations for these factors (by the amount shown in the top two rows of each matrix, in the last three columns). The off–diagonal terms determine the bias that would be introduced if univariate correction methods (i.e. just using the diagonal terms to adjust the associations) were employed in these multivariate settings.

As an example, the first matrix (which illustrates these effects for total cholesterol and

systolic blood pressure) indicates that in order to calculate the true association between total cholesterol 20 years after baseline and disease risk around that time, one should multiply the baseline association by 1.84 (equivalently, divide by 0.54) and add 0.04 times the baseline association for systolic blood pressure (recall that this association corresponds to a 20 mmHg change in SBP). Similarly, the true association for systolic pressure is obtained from the baseline association by multiplying it by 2.13 and adding 0.3 times the baseline association for total cholesterol. In this example, the true associations for age and cigarette smoking are virtually the same as the baseline associations, as they are modified very little through their relationships with total cholesterol and systolic pressure (since the modifying factors are all relatively small (between –0.05 and 0.08)).

Overall, it can be seen that the associations for age and cigarette smoking are modified quite substantially by their relations with diastolic pressure (second row of $\mathbf{\Lambda}_2^{-1}$ and $\mathbf{\Lambda}_4^{-1}$), much less so by their relations with systolic pressure (second row of $\mathbf{\Lambda}_1^{-1}$ and $\mathbf{\Lambda}_3^{-1}$), virtually not at all by their relations with HDL cholesterol (first row of $\mathbf{\Lambda}_3^{-1}$ and $\mathbf{\Lambda}_4^{-1}$), and only marginally by their relations with total cholesterol (first row of $\mathbf{\Lambda}_1^{-1}$ and $\mathbf{\Lambda}_2^{-1}$).

5.5 Discussion

5.5.1 Interpretation of findings

These results suggest that for continuous risk factors, the use of baseline levels of physical, biochemical and haemostatic measurements in univariate analyses leads to marked underestimation of their associations with disease risk, and that the extent of this underestimation increases with duration of follow–up. Long–term associations with CHD (say those occurring 15–25 years after baseline) would be underestimated by about one half for systolic blood pressure and blood lipids (total cholesterol, LDL cholesterol and triglyceride) and three quarters for diastolic blood pressure, while associations between glucose tolerance or insulin resistance and CHD would be underestimated by two thirds or more if based on single baseline measures of insulin and glucose. In contrast, regression dilution effects for HDL cholesterol and the ratio of total to HDL cholesterol were less marked, so that the use of a baseline HDL measurement to estimate associations with disease risk 20 years later would only result in underestimation by about one quarter. Even in the short–term however, say over a period of a few weeks or months, associations with usual

risk exposure levels would still be underestimated by between 10 and 20%, and for blood pressure, this figure is likely to be nearer 30%.

For the categorical 'lifestyle' risk factors, substantial within–person variation was also evident. Nearly one half of all men had their physical activity level or alcohol intake group defined differently at baseline from that defined according to usual reported levels (taking into account follow–up questionnaires). This reflects the general decline in physical activity and alcohol intake observed with increasing age, but also reflects the tendency for individuals to have, upon re–measurement, exposure levels that are 'less extreme' than observed at baseline (as happens for continuous variables). While a greater degree of agreement was observed for cigarette smoking, a combination of secular trends in smoking habits and true measurement errors in the ascertainment of smoking exposure together led to a substantial overestimation of the proportion of men estimated to be regular heavy smokers. This could have implications for the estimation of relationships between disease risk and true levels of cigarette smoking.

In analyses relating blood lipids, blood pressure, age and cigarette smoking simultaneously to disease risk, it would seem that the use of total cholesterol together with systolic blood pressure would result in the least residual bias from using univariate correction methods to adjust for within–person variation in a multivariate setting.

5.5.2 Validity of analyses: estimating the RDR

The 'regression based' methods used in the analyses presented in this chapter to estimate the regression dilution ratio can be used to calculate a valid correction factor when the distributions of the baseline and follow–up measurements are different (as may be expected in an ageing cohort), and it is for this reason that they have been preferred to the other approaches, particularly the correlation approaches, described in chapter 3.[20] One potential problem in the estimates of regression dilution presented in this chapter however, is that they are based on different follow–up periods on different individuals at somewhat different ages. However, given that the estimates of the RDR were fairly similar between individuals of different ages (see Figures 5.6 to 5.11), it is likely that this potential source of bias is not of paramount importance. The estimates were also generally found to be independent of geographical area or social class, and may therefore be generalisable to other populations. However, one factor that may affect estimates of regression dilution bias, and

hence the extent that it may be appropriate to use the estimates in other settings, is the proportion of the population with evidence of CHD at baseline. Men who developed CHD during the study were certainly observed to experience a greater degree of within–person variation in their risk factors than men who remained disease free, though it is hypothesised that much of this variation would occur after the CHD event had occurred. It may therefore be prudent to ensure that the relevance of the RDR estimate to the population for which it is to be applied, is considered before correction for regression dilution bias is performed. In general however, our estimates of the regression dilution ratio are similar to those from many other prospective studies of cardiovascular disease, the findings of which are now described.

5.5.3 Comparison with other studies

Blood pressure and total cholesterol

Many previous cardiovascular studies have explored the effects of regression dilution bias for total cholesterol and blood pressure, two of the strongest CHD risk factors. A summary of the estimates from some of these studies is shown in Figure 5.12.[17;19;517–522] One of the first studies to assess these effects was an analysis of nine prospective studies, which aimed to estimate the associations between usual levels of diastolic blood pressure and the risks of CHD and stroke.[17] The authors used repeated DBP measurements from several studies including the Framingham study to estimate the effects of regression dilution bias in DBP over a four year period and, from these data, estimated that the true associations would be about 60% greater than that estimated from baseline measurements of DBP (corresponding to a regression dilution ratio for diastolic pressure of 0.62). Since this analysis, numerous subsequent studies have published estimates of regression dilution ratios for blood pressure and total cholesterol. In ten cohorts of Japanese and Chinese individuals, repeat measurements of diastolic blood pressure and total blood cholesterol were taken over an average five year interval.[517] Over this period, regression dilution ratio estimates of 0.48 for diastolic pressure and 0.52 for total cholesterol were observed – the figure for total cholesterol being somewhat lower than our estimate taken over a four year period. In the Israel Ischaemic Heart Disease Project, regression dilution ratios for systolic and diastolic pressure over a two year period were 0.65 and 0.54 respectively,[518] while in the Finrisk Haemostasis study, estimates of 0.70 and 0.61 for these measures were

obtained over a three year interval.[519] Among men in the Renfrew/Paisley study, RDR estimates of 0.56 were observed for both systolic and diastolic blood pressure over a four year interval; for men in the Collaborative study, these estimates were approximately 0.63 over five years.[520]

Though numerous studies have provided estimates of the regression dilution ratio over short periods of follow up, fewer studies have been able to provide long–term estimates (>10 years). One which has is the Framingham study, where after 16 and 26 years of follow–up RDR estimates of 0.52 and 0.34 were observed for systolic pressure, 0.38 and 0.26 for diastolic pressure, and 0.52 and 0.43 for total cholesterol.[19] Similarly, in the Whitehall I study, RDR estimates of 0.32, 0.29 and 0.28 were calculated for SBP, DBP and total cholesterol over a 26 year interval.[19] In a study of 6,137 middle–aged men of Japanese descent in the Honolulu Heart Program, repeated total cholesterol measurements were available over a 16 year interval.[521] The authors found that the crude baseline association between serum total cholesterol and CHD death underestimated the true association by 40% (RDR=0.6). Correspondingly, their corrected association was 67% greater (since $1/0.6 = 1.67$) than their uncorrected association (though in their paper they quoted a 22% increase, this is misleading as it should have been calculated on the log scale). Resurveys of study participants after 30 years of follow–up for men and women in the Glostrup '1914–cohort' and the Framingham study estimated regression dilution ratios as low as 0.27 for systolic pressure and 0.14 for diastolic pressure.[522] It can be seen from Figure 5.12 that our own long–term estimates taken over 20 years are fairly consistent with these other studies.

Our observation that the effects of regression dilution were more marked amongst individuals developing CHD over the interval period (Figures 5.6 to 5.11) is also consistent with data from three large studies of patients with cardiovascular disease: the UK–TIA trial,[524] the Dutch TIA trial[525] and the European Carotid Surgery Trial.[526] These studies included regular repeated blood pressure measurements over intervals of between 5 months and 3 years, and found that even after only 4–5 months, use of baseline measures in analyses would result in usual associations with disease risk over that period being underestimated by 50% or more.[527]

Novel risk factors

Since the mechanisms behind regression dilution bias applies to any setting in which the primary interest is to estimate the prospective relationship between usual risk exposure levels and the risk of a particular disease occurring, interest has naturally shifted in recent times to the effects of regression dilution bias for new risk factors for CHD. For instance, a recent study of repeated homocysteine measurements measured in the Rotterdam Scan study, the Hordalan study, the Framingham study and the United Kingdom Prospective Diabetes Study revealed regression dilution ratio estimates for homocysteine of 0.83 at 2 years, 0.71 at 6 years and 0.53 at 12 years.[528] In their analyses, the authors used the quintile method to estimate the RDR (see section 3.5.3); when the 4–year repeated data were re–analysed using this method, an RDR estimate for homocysteine of 0.76 was obtained – highly consistent with the results from these four other studies. In the Cholesterol and Recurrent Events (CARE) trial of the effects of pravastatin in patients with a recent myocardial infarction,[79] C–reactive protein was measured at baseline and after five years of follow–up in a random sample of 472 participants who remained free of recurrent coronary events during follow–up. The correlation between the baseline and follow–up measurements (which may also be used as an estimate of the regression dilution ratio) for these individuals was 0.60, almost exactly the same as our estimate taken over a similar period.[529] Recent results from the Reykjavik prospective study provide an estimate for C–reactive protein of 0.59 over a 12–year period.[346] Despite the apparent inconsistency, reference to Figure 5.4 reveals that this estimate is also in general agreement with the estimates in this chapter. For haemostatic factors, fibrin D–dimer and von Willebrand factor were measured in 1,009 subjects over a 5–year period in the Edinburgh Artery Study. The correlation between fibrin D–dimer levels measured at baseline and after 5 years was 0.51,[356] somewhat lower than our estimate. However, for von Willebrand factor, the correlation was 0.63,[358] which although lower is not inconsistent with our 4–year estimate of 0.71. Fibrinogen was measured over a 3–year period in a sample of 473 men and women in the Finrisk Haemostasis study – a correlation of 0.72 was observed between these measurements.[519]

5.5.4 Is correction always appropriate?

Correction for regression dilution bias is a useful tool for estimating associations between usual or average exposure levels and a particular outcome of interest and, as we have demonstrated, failure to take these effects into account can lead to true associations being greatly underestimated. However, it should be stressed that in situations where the aetiological relationship is not with the underlying usual value of the predictor, adjustment for regression dilution bias may not be appropriate.[530] For instance, suppose that recent levels of an exposure, or perhaps peaks in an exposure, were more important predictors of a particular outcome than the usual level. Attempting to estimate the association with the 'usual" level in these situations would not necessarily be appropriate. In the Intersalt study, for example, the association between sodium excretion and blood pressure was corrected for regression dilution bias.[531] The validity of performing such a correction was subsequently strongly debated, both because of the relevance of recent (and not usual) levels in determining this relationship and also because of the problems caused by the potential correlation between the errors in measurement of the predictor and the outcome.[530;532] In the context of quantifying the long–term observational relationships between coronary risk factors and coronary heart disease risk however (where the disease outcome is measured with precision), it would seem perfectly reasonable to assume that risk is determined by average risk exposures during adulthood (or at least average risk exposure levels over, say, the previous decade), and hence it would be appropriate to use the estimates of regression dilution bias presented in this chapter to correct estimated exposure–disease relationships. Where it may not necessarily be appropriate to correct for regression dilution bias however, is in the use of risk prediction equations (e.g. the Framingham equations),[50;410] because these equations are commonly derived from the same information that will be available to the clinician attempting to use them (i.e. risk factor levels measured at a single point in time).

5.5.5 Conclusions: extent of within–person variation

In this chapter the extent of within–person variation in both continuous and categorical risk factors for CHD has been estimated for men in the British Regional Heart Study. For continuous risk factors, this variation has been quantified in terms of the extent that univariate relationships with disease risk over particular periods of follow–up would be

underestimated due to the use of baseline measures in analyses (the extent of 'regression dilution bias'). These effects have been shown to be substantial. For categorical 'lifestyle' risk factors, estimated average exposures to these factors over the 'risk period' have been calculated through a combination of the baseline and follow–up questionnaires. For these factors, the extent of within–person variation (in terms of the degree of misclassification) was also found to be considerable, particularly amongst certain high–risk exposure groups (e.g. heavy smokers, heavy drinkers). In a multivariate analysis setting, correlations between different risk factors can lead to under or overestimation of risk associations, even for risk factors measured precisely. Under such circumstances, the appropriateness of performing univariate correction methods may be evaluated by calculating a "modifying matrix" $\mathbf{\Lambda}^{-1}$.

Table 5.1: Selected characteristics over periods of 1 week, 4, 16 and 20 years. For each measure, men that attended only the first screening in the pair are presented alongside those that attended both screenings. Data indicate mean (SD) unless otherwise indicated.

	1 week period		4 year period		16 year period		20 year period	
	Initial	FU	Initial	FU	Initial	FU	Initial	FU
Year of screening	2000	2000	1996	2000	1980	1996	1980	2000
Number of men	112	112	297	297	425	425	4252	4252
Age (years)	70.0 (5.2)	70.0 (5.2)	65.0 (5.5)	69.0 (5.5)	49.6 (5.7)	65.6 (5.7)	48.9 (5.5)	68.9 (5.5)
Total cholesterol (mmol/L)	5.7 (1.1)	5.5 (1.0)	5.9 (1.1)	5.9 (1.1)	6.4 (1.0)	5.9 (1.0)	6.3 (1.0)	6.0 (1.1)
LDL cholesterol (mmol/L)	3.8 (1.0)	3.7 (0.9)	3.6 (1.0)	3.8 (1.0)	4.3 (0.9)	3.7 (0.9)	4.2 (1.0)	3.9 (1.0)
HDL cholesterol (mmol/L) †	1.12 (0.90–1.30)	1.07 (0.90–1.20)	1.24 (1.00–1.40)	1.27 (1.10–1.50)	1.14 (0.99–1.30)	1.23 (1.00–1.40)	1.12 (0.97–1.29)	1.28 (1.10–1.50)
Triglyceride (mmol/L) †	1.4 (1.1–1.8)	1.4 (1.0–1.8)	2.0 (1.4–2.7)	1.6 (1.1–2.2)	1.8 (1.2–2.7)	2.0 (1.4–2.9)	1.7 (1.2–2.5)	1.6 (1.2–2.2)
SBP (mm Hg)	149 (24)	140 (21)	149 (25)	151 (25)	148 (19)	149 (25)	143 (20)	149 (24)
DBP (mm Hg)	85.8 (11.7)	83.0 (10.4)	85.0 (11.2)	86.4 (10.8)	83.5 (13.1)	84.3 (11.0)	81.4 (12.8)	85.1 (11.2)
Glucose (mmol/L) †	6.1 (5.5–6.3)	6.0 (5.5–6.4)	5.7 (5.1–5.9)	6.4 (5.6–6.4)	5.4 (5.0–5.9)	5.8 (5.1–6.2)	5.5 (4.9–5.9)	5.9 (5.3–6.1)
Insulin (mU/L) †	8.6 (6.1–11.7)	8.5 (5.9–12.0)	10.7 (5.6–19.1)	8.4 (5.4–11.8)	12.7 (7.4–21.2)	10.9 (5.6–20.6)	12.5 (7.3–20.5)	8.7 (5.7–12.4)

†Geometric mean (interquartile range)

Table 5.2: Estimates of the regression dilution ratio (RDR) by duration of follow-up for the blood lipids, blood pressure, body mass index, insulin and glucose.

Risk exposure	RDR (95% CI)			
	One week	4 years	16 years	20 years
Total cholesterol	0.87 (0.80,0.94)	0.70 (0.61,0.80)	0.59 (0.50,0.67)	0.53 (0.50,0.56)
HDL cholesterol §	1.00 (0.93,1.06)	0.87 (0.79,0.95)	0.68 (0.60,0.76)	0.74 (0.71,0.77)
LDL cholesterol	0.86 (0.79,0.94)	0.61 (0.51,0.71)	0.49 (0.40,0.59)	0.48 (0.45,0.51)
Triglyceride §	0.86 (0.73,0.98)	0.66 (0.58,0.75)	0.47 (0.39,0.56)	0.43 (0.41,0.46)
Total:HDL §	0.95 (0.89,1.00)	0.77 (0.69,0.86)	0.65 (0.57,0.74)	0.64 (0.61,0.67)
Body mass index	1.02 (1.00,1.03)	0.91 (0.84,0.97)	0.94 (0.86,1.02)	0.93 (0.91,0.96)
SBP	0.69 (0.58,0.80)	0.61 (0.52,0.70)	0.66 (0.55,0.77)	0.49 (0.45,0.52)
DBP	0.72 (0.62,0.83)	0.51 (0.41,0.60)	0.27 (0.19,0.35)	0.24 (0.22,0.27)
Mid blood pressure	0.70 (0.60,0.80)	0.59 (0.50,0.68)	0.48 (0.38,0.57)	0.40 (0.37,0.43)
Mean arterial pressure	0.71 (0.60,0.81)	0.57 (0.48,0.67)	0.40 (0.31,0.49)	0.35 (0.32,0.38)
Glucose §	0.83 (0.71,0.96)	0.52 (0.40,0.63)	0.29 (0.19,0.39)	0.33 (0.30,0.37)
Insulin §	0.87 (0.72,1.02)	0.34 (0.26,0.42)	0.30 (0.20,0.41)	0.27 (0.24,0.30)

§Analysed on the log scale; CI = confidence interval; SBP = systolic blood pressure; DBP = diastolic blood pressure; Mid blood pressure = (SBP + DBP)/2; Mean arterial pressure = (1/3) SBP + (2/3) DBP.

Table 5.3: Estimates of the regression dilution ratio (RDR) for the novel coronary risk factors by duration of follow–up

Risk exposure	RDR (95% CI)			
	One week	4 years	16 years	20 years
Homocysteine §	0.87 (0.81,0.93)	0.81 (0.73,0.90)	0.49 (0.32,0.67)	0.41 (0.31,0.50)
C–reactive protein §	0.74 (0.61,0.87)	0.61 (0.51,0.72)	0.50 (0.27,0.72)	0.42 (0.35,0.49)
Fibrinogen	0.84 (0.75,0.93)	0.52 (0.41,0.64)	–	–
Factor VII	0.88 (0.80,0.96)	0.75 (0.65,0.85)	–	–
vWF §	0.83 (0.73,0.92)	0.71 (0.64,0.78)	0.67 (0.49,0.85)	0.51 (0.46,0.56)
t–Pa §	0.71 (0.58,0.83)	0.68 (0.58,0.79)	0.33 (0.12,0.53)	0.34 (0.28,0.39)
D–dimer §	0.85 (0.78,0.93)	0.76 (0.67,0.86)	0.20 (–0.04,0.43)	0.17 (0.11,0.24)

§Analysed on the log scale; CI = confidence interval; SBP = systolic blood pressure; DBP = diastolic blood pressure; vWF = Von Willebrand Factor

Table 5.4: Responses to each of the study questionnaires regarding cigarette smoking, physical activity and alcohol intake. Figures indicate column percentages.

	Study Questionnaire				
	Baseline (n=7735)	Q5 (n=7275)	Q92 (n=5925)	Q96 (n=5263)	Q20 (n=4252)
Cigarette smoking					
Never smoked	23.6	24.1	26.1	27.2	29.0
Ex-smoker	35.2	43.6	54.5	57.5	58.3
Current (1–20 a day)	26.2	23.7	16.0	12.9	11.0
Current (21–39 a day)	11.0	6.4	2.6	2.1	1.3
Current (≥40 a day)	4.1	2.3	0.9	0.4	0.5
Physical activity					
None	9.0	–	7.7	15.1	11.5
Occasional	30.7	–	26.7	26.6	23.4
Light	23.1	–	22.9	19.6	18.7
Moderate	15.8	–	15.9	13.0	14.4
Moderately vigorous	14.7	–	15.5	14.7	16.8
Vigorous	6.7	–	11.3	11.1	15.2
Alcohol intake					
None	6.0	9.8	17.4	16.5	10.4
Occasional	23.9	29.7	23.8	21.8	27.0
Light	32.9	37.1	41.0	44.1	43.9
Moderate	26.4	19.3	14.3	14.3	15.7
Heavy	10.8	4.1	3.5	3.3	3.0

Table 5.5: Comparison of cigarette smoking status measured at baseline with "average" cigarette smoking status over the follow-up period – no. (row %)

	"Average" exposure to cigarette smoking during the study						
Baseline smoking status	Never	Ex	New*	Current (1–20)	Current (21–39)	Current (≥40)	Total
Never	1802 (99)	–	17 (1)	–	–	–	1819 (100)
Ex–smoker	–	2525 (93)	190 (7)	–	–	–	2715 (100)
Current (1–20)	–	–	–	1936 (96)	81 (4)	6 (0)	2023 (100)
Current (21–39)	–	–	–	447 (53)	393 (46)	6 (1)	846 (100)
Current (≥40)	–	–	–	106 (34)	127 (40)	83 (26)	316 (100)
Total	1802 (23)	2525 (33)	207 (3)	2489 (32)	601 (8)	95 (1)	7719 (100)

* New or recurrent cigarette smoker

Table 5.6: Comparison of physical activity level measured at baseline with "average" physical activity level over the follow-up period – no. (row %)

	"Average" physical activity level during the study						
Baseline physical activity level	None	Occasional	Light	Moderate	Moderately vigorous	Vigorous	Total
None	389 (57)	227 (33)	52 (8)	18 (3)	–	–	686 (100)
Occasional	128 (5)	1470 (63)	504 (21)	216 (9)	27 (1)	–	2345 (100)
Light	–	378 (21)	1062 (60)	252 (14)	69 (4)	–	1761 (100)
Moderate	–	55 (5)	366 (30)	576 (48)	208 (17)	–	1205 (100)
Mod. vigorous	–	8 (1)	145 (13)	282 (25)	586 (52)	99 (9)	1120 (100)
Vigorous	–	–	23 (4)	82 (16)	140 (27)	268 (52)	513 (100)
Total	517 (9)	2138 (31)	2152 (23)	1426 (16)	1030 (15)	367 (7)	7630 (100)

Table 5.7: Comparison of alcohol intake measured at baseline with "average" alcohol intake over the follow-up period – no. (row %)

	"Average" alcohol intake during the study					
Baseline alcohol intake	None	Occasional	Light	Moderate	Heavy	Total
None	371 (80)	80 (17)	12 (3)	3 (1)	–	466 (100)
Occasional	327 (18)	1147 (62)	357 (19)	14 (1)	–	1845 (100)
Light	46 (2)	655 (26)	1654 (65)	188 (7)	1 (0)	2544 (100)
Moderate	10 (0)	195 (10)	975 (48)	806 (39)	56 (3)	2042 (100)
Heavy	1 (0)	24 (3)	143 (17)	430 (52)	234 (28)	832 (100)
Total	755 (10)	2101 (27)	3141 (41)	1441 (19)	291 (4)	7729 (100)

Figure 5.1: Four hypothetical examples illustrating the method used to calculate average exposures over the follow–up period from the study questionnaires (Q0, Q5, Q92, Q96 and Q20). In each case assume the man was enrolled into the study in November 1980 so that the intervals between successive questionnaires are 5 years, 7 years, 4 years and 4 years respectively.

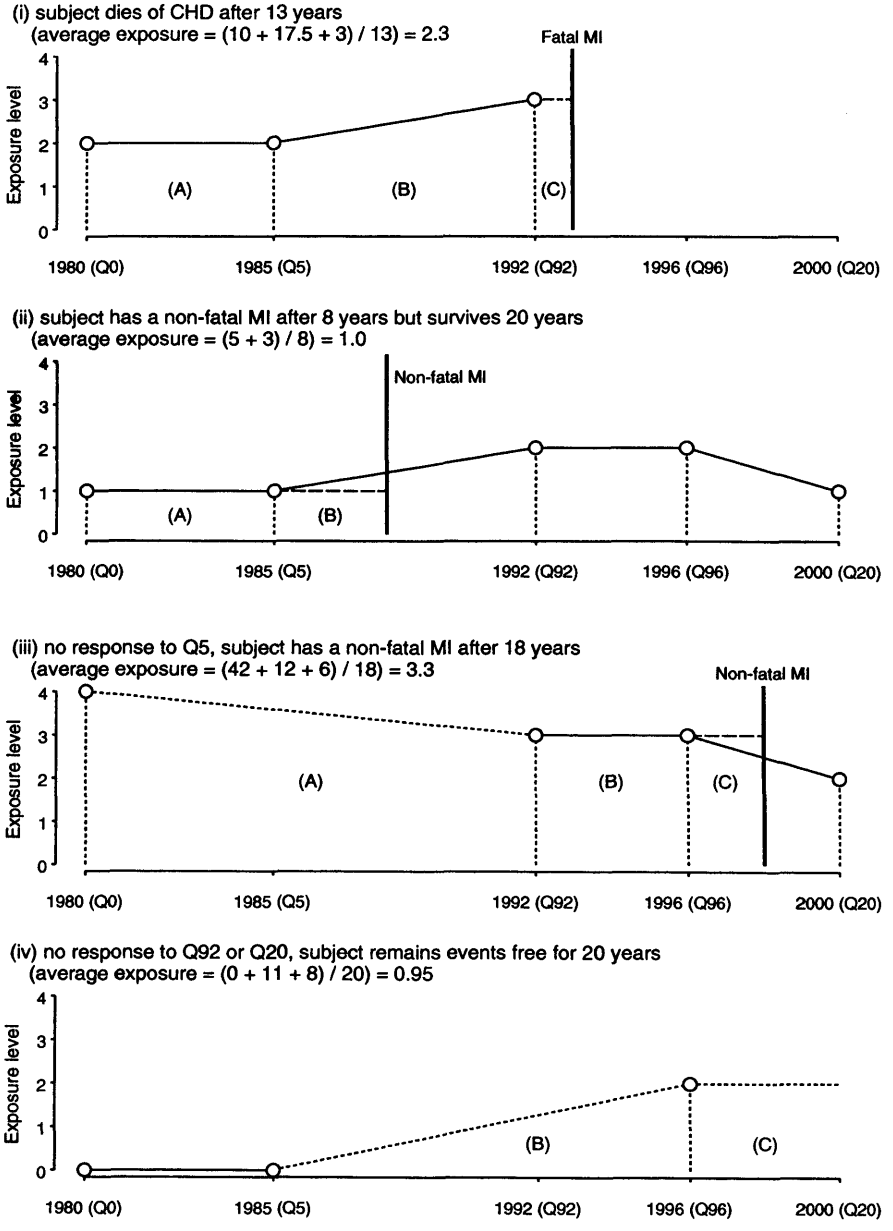


Figure 5.2: Regression dilution ratio estimates (with 95% CIs) for blood lipid measurements and body mass index by length of interval period.

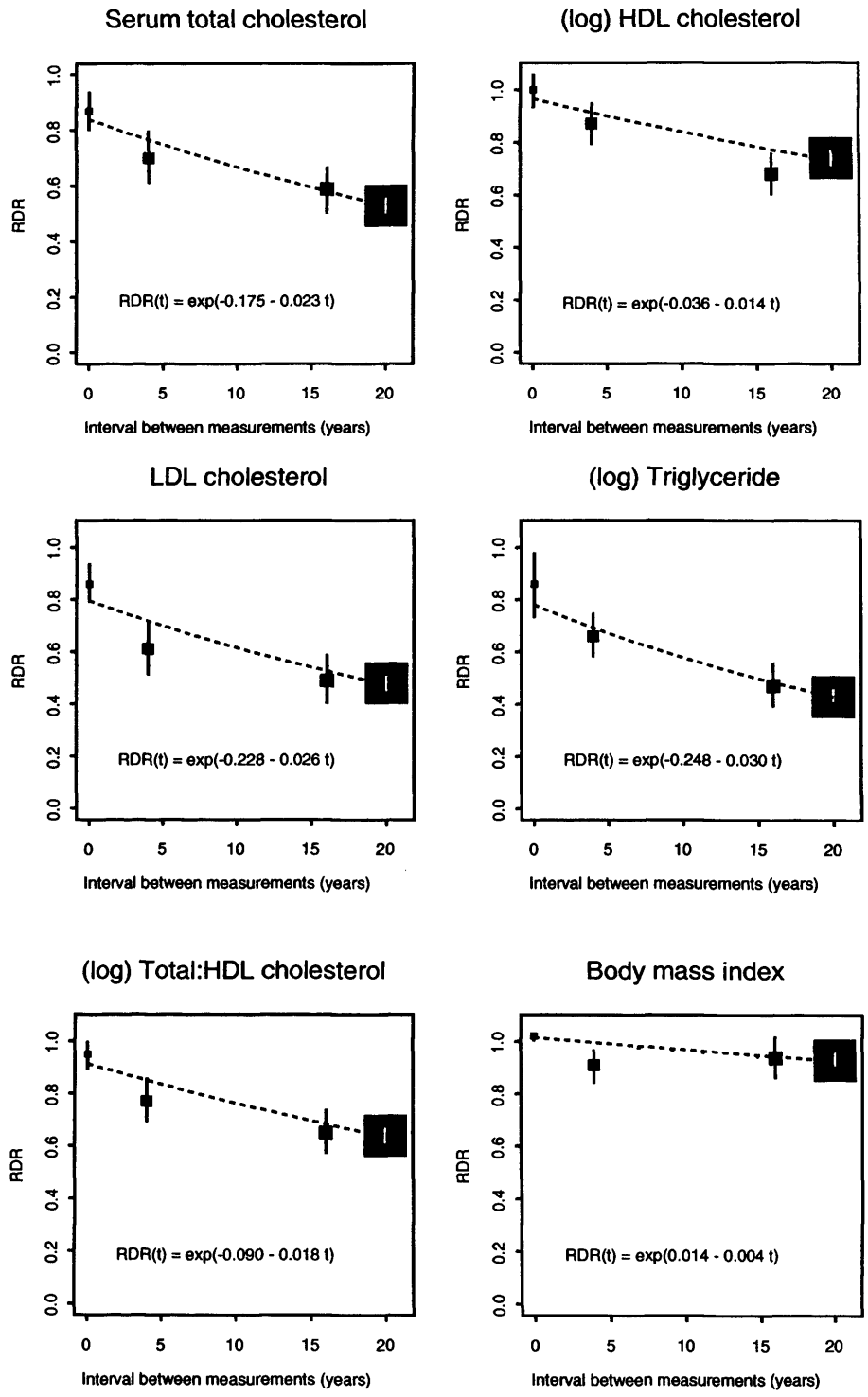


Figure 5.3: Regression dilution ratio estimates (with 95% CIs) for blood pressure, insulin and glucose by length of interval period.

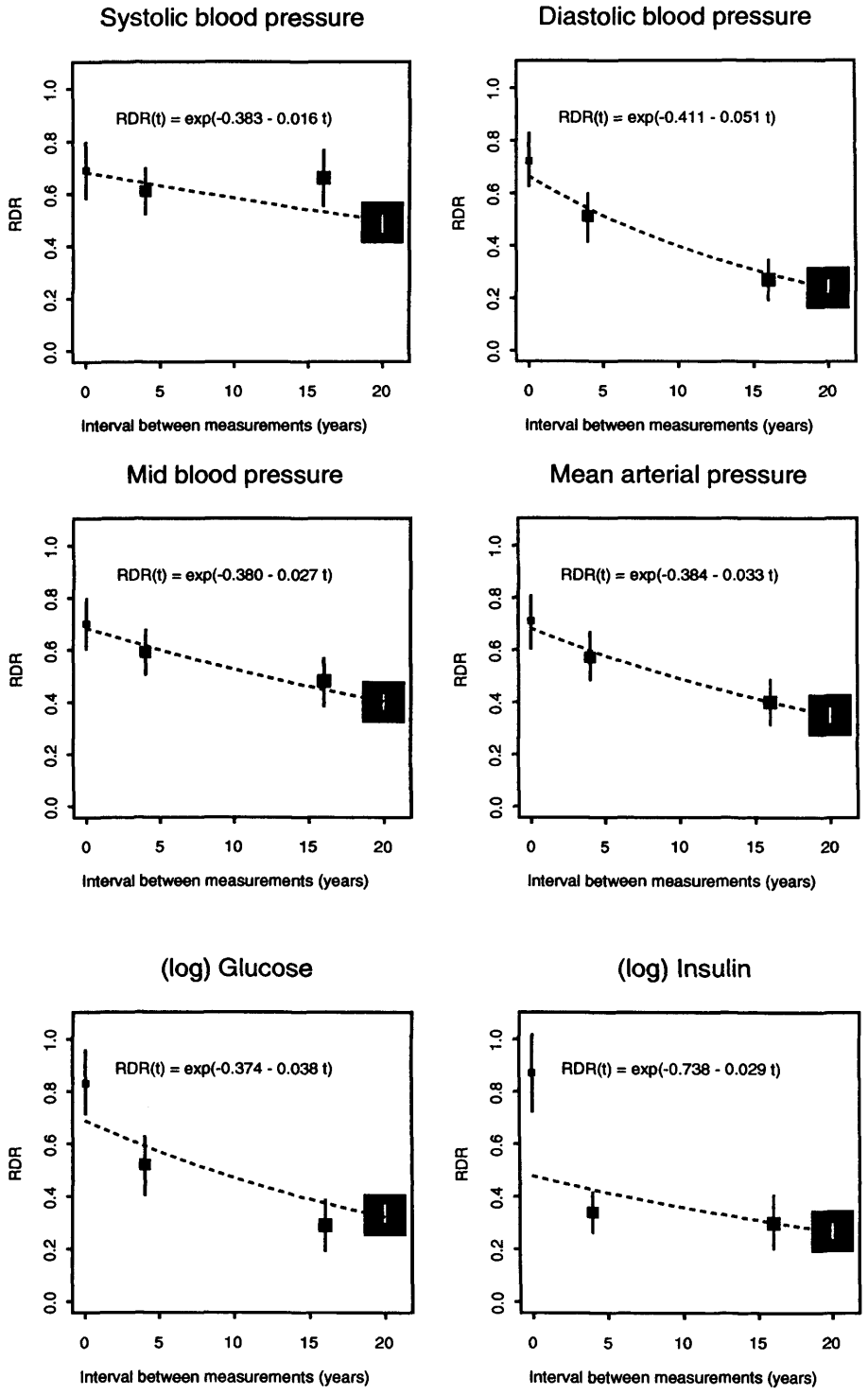


Figure 5.4: Regression dilution ratio estimates (with 95% CIs) for homocysteine and C-reactive protein by length of interval period.

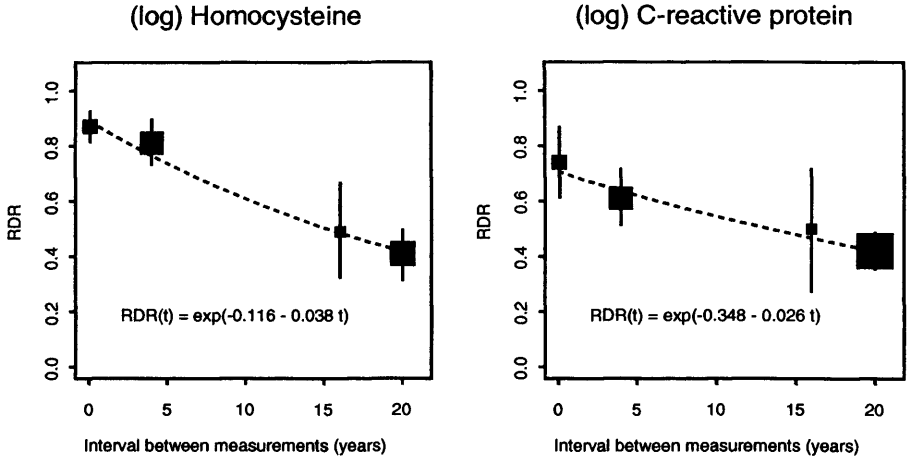


Figure 5.5: Regression dilution ratio estimates (with 95% CIs) for the haemostatic factors by length of interval period.

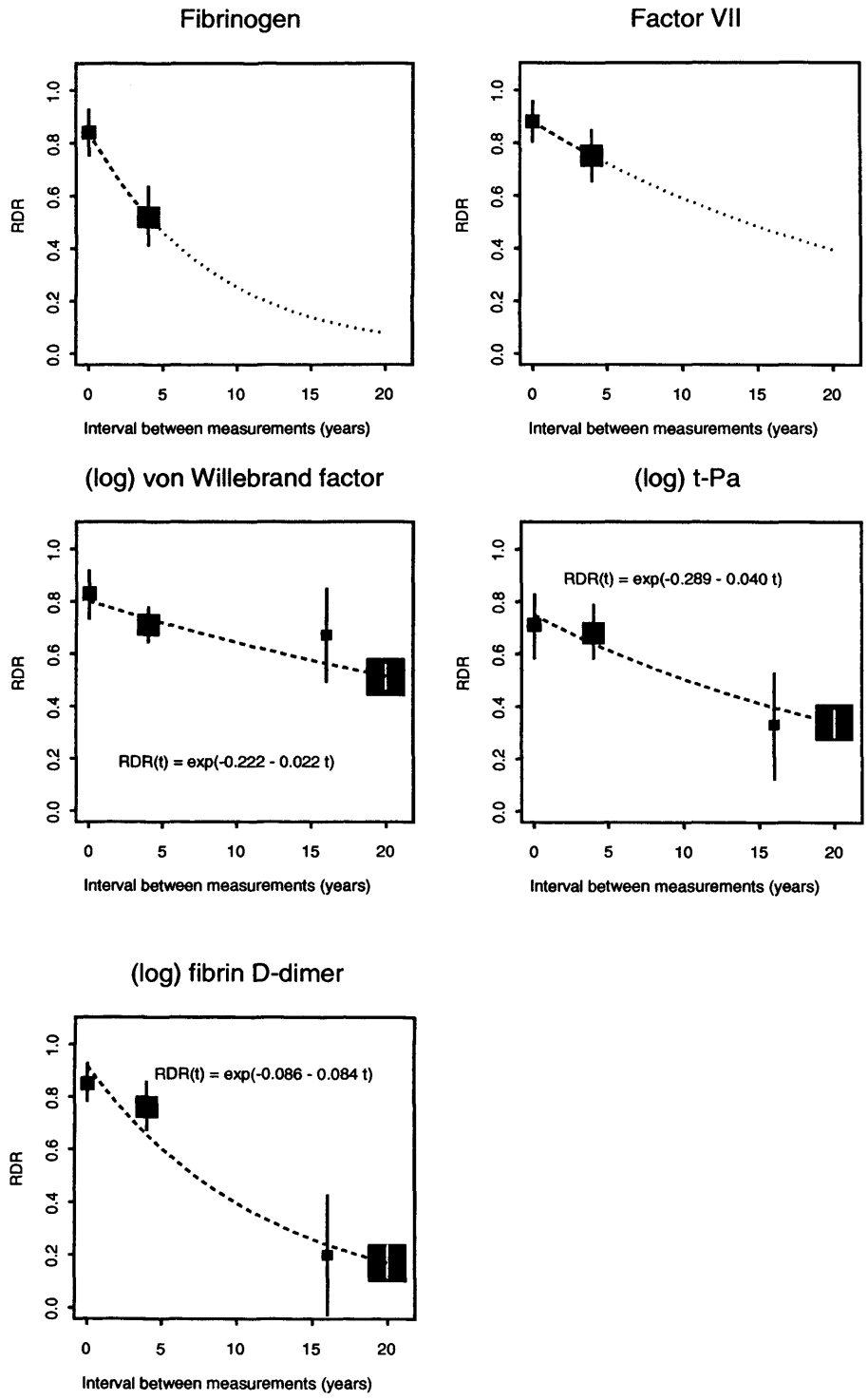


Figure 5.6: Estimates of the RDR for total cholesterol over 20 years by age, history of CHD, social class and geographical location.

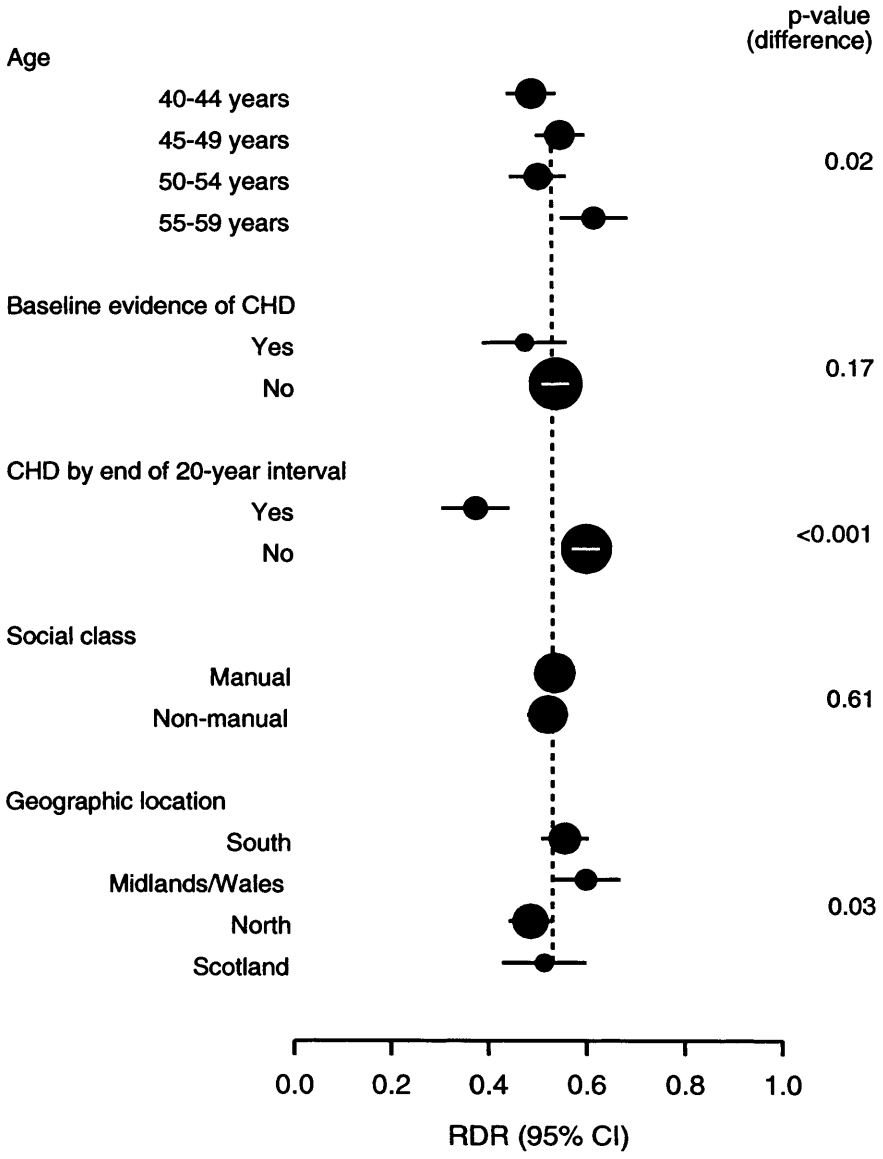


Figure 5.7: Estimates of the RDR for HDL cholesterol over 20 years by age, history of CHD, social class and geographical location.

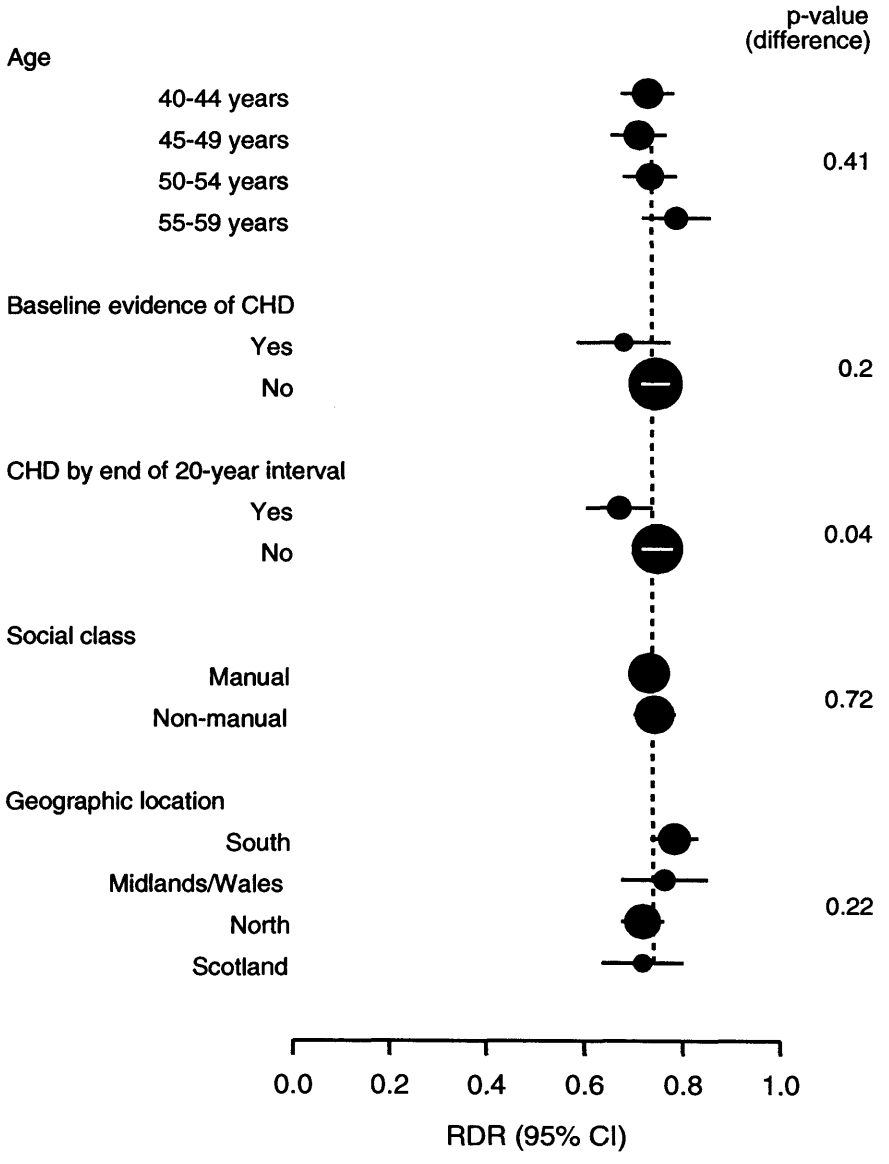


Figure 5.8: Estimates of the RDR for LDL cholesterol over 20 years by age, history of CHD, social class and geographical location.

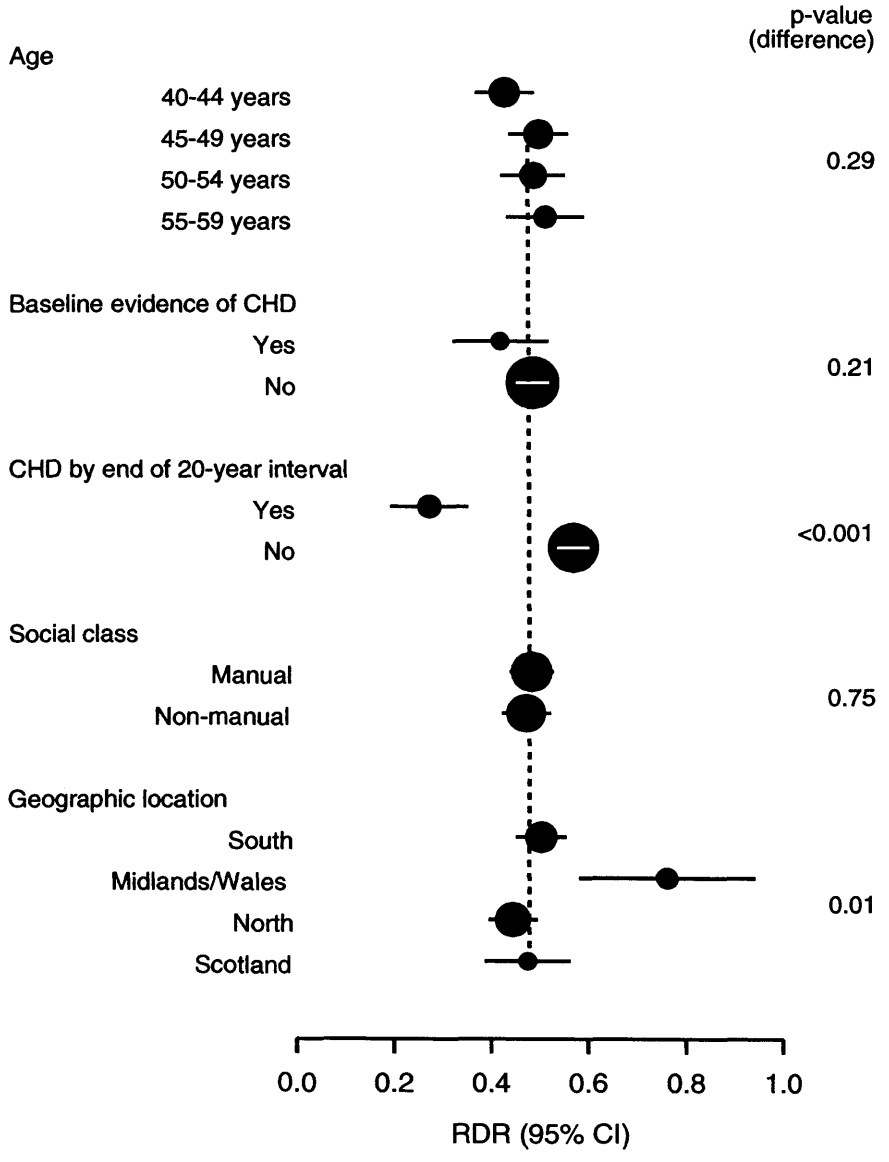


Figure 5.9: Estimates of the RDR for the ratio of total to HDL cholesterol over 20 years by age, history of CHD, social class and geographical location.

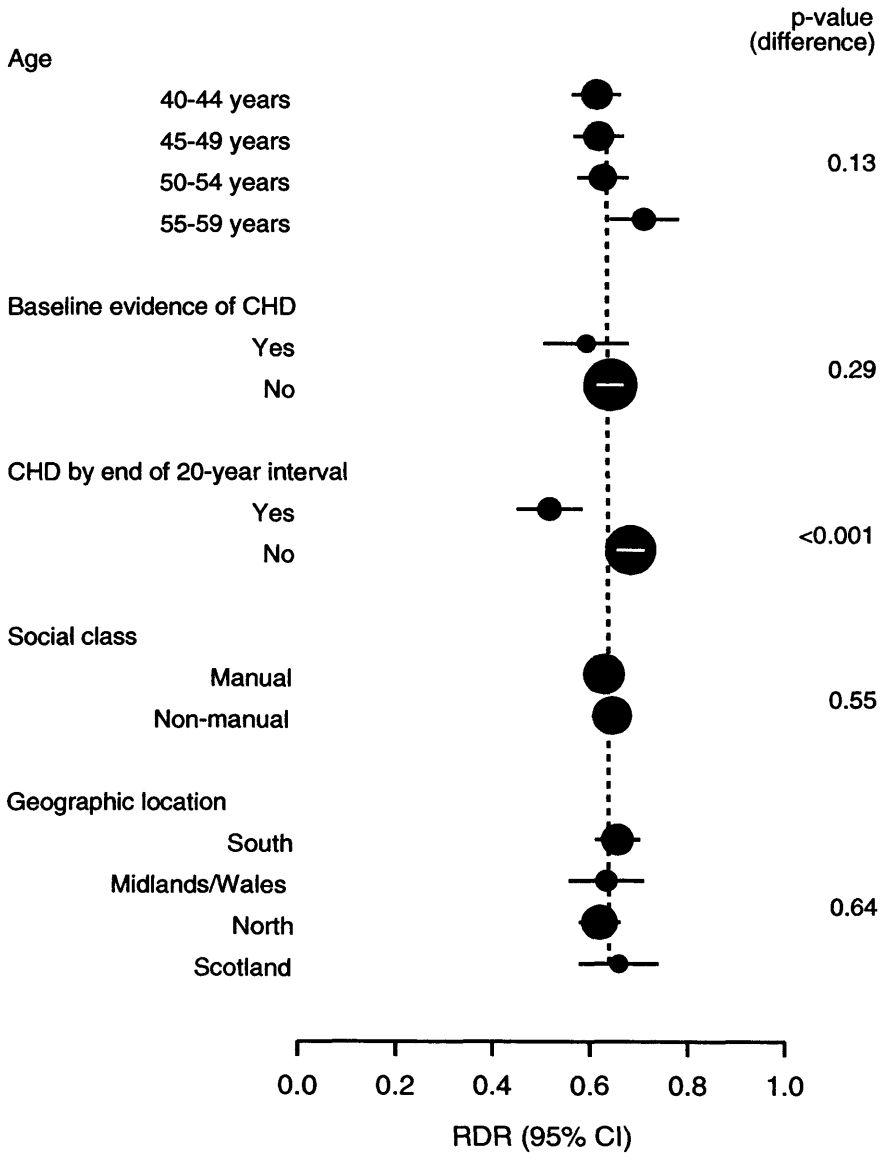


Figure 5.10: Estimates of the RDR for systolic blood pressure over 20 years by age, history of CHD, social class and geographical location.

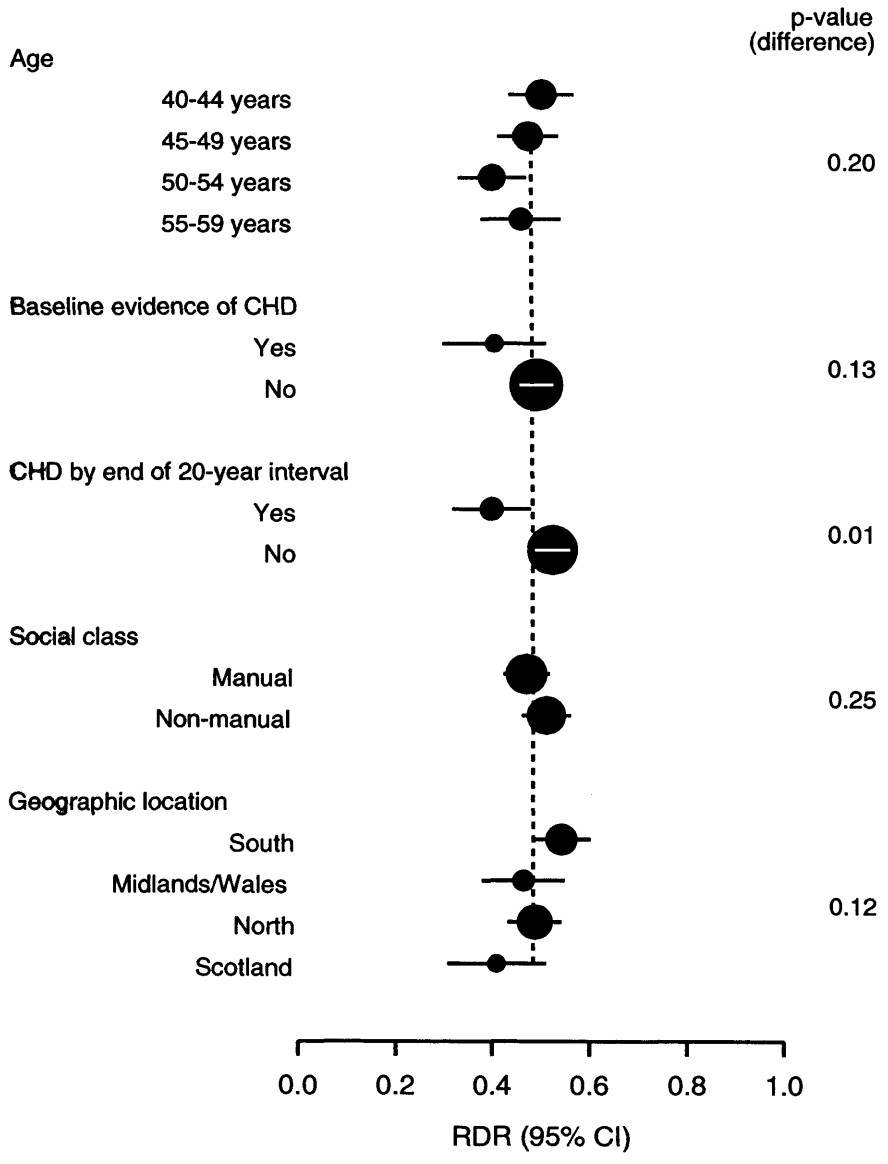


Figure 5.11: Estimates of the RDR for diastolic blood pressure over 20 years by age, history of CHD, social class and geographical location.

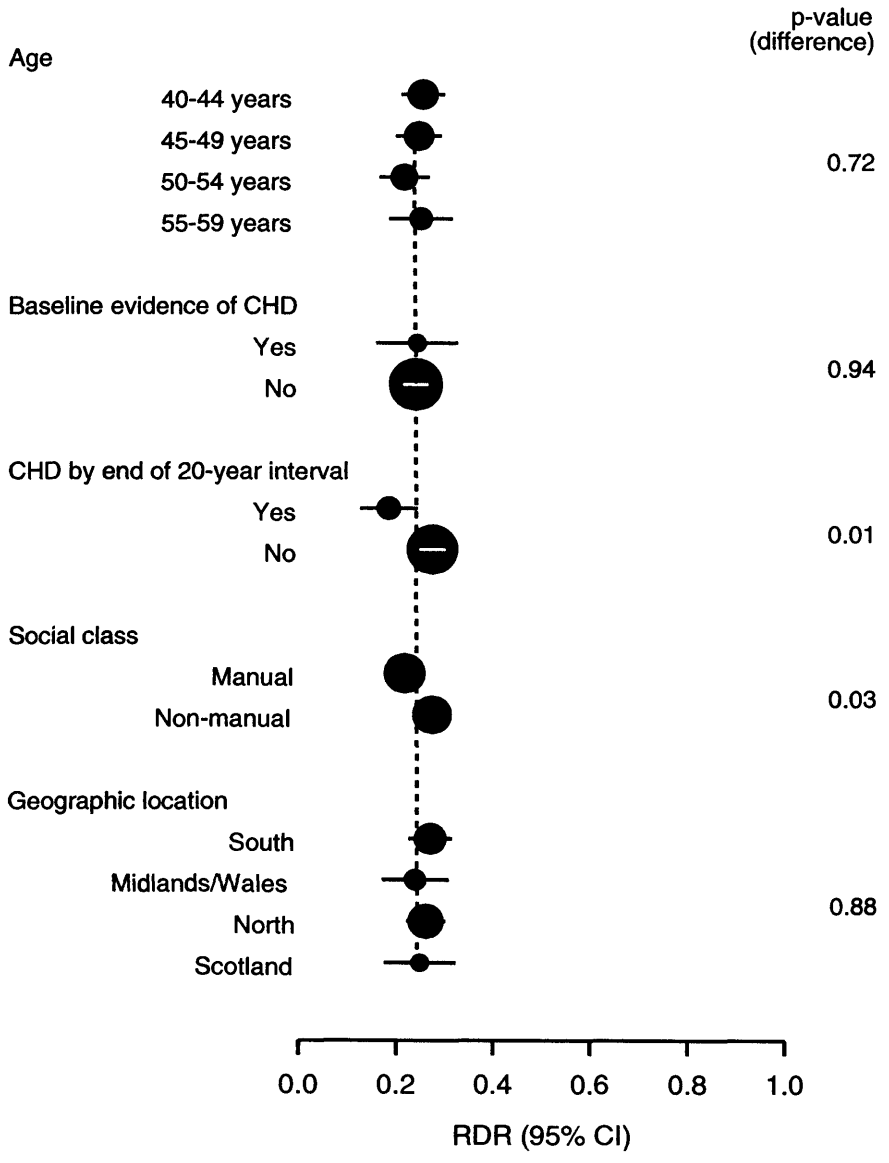
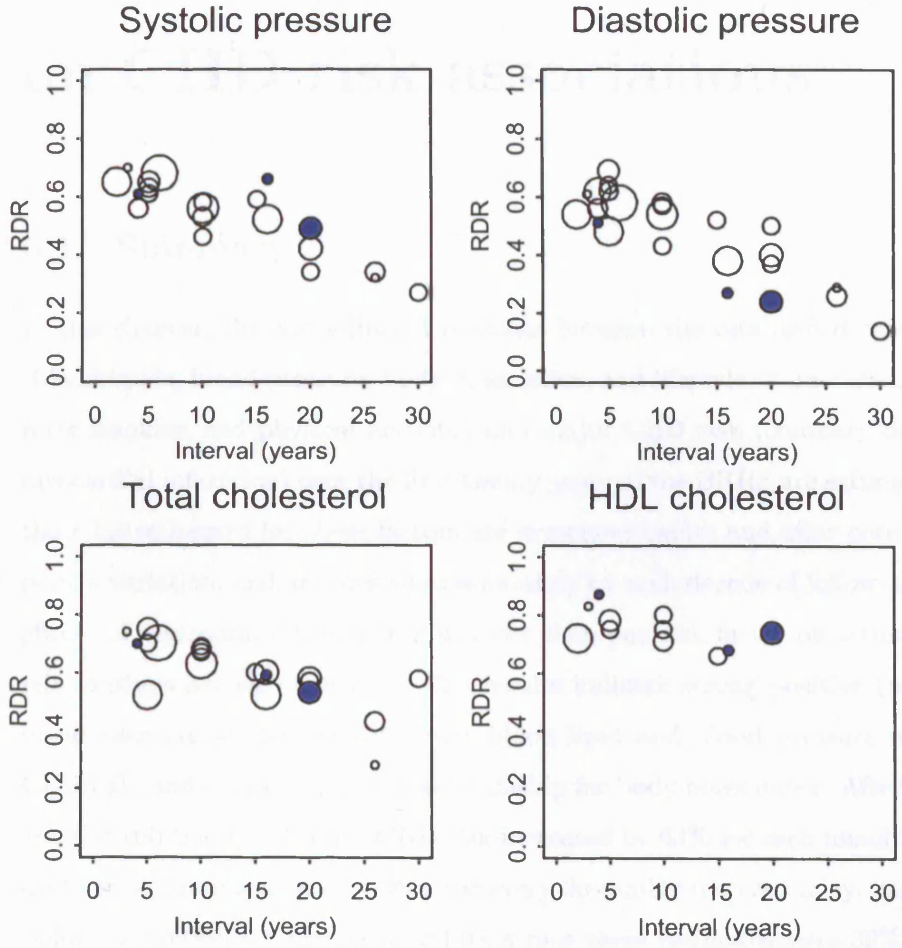


Figure 5.12: Estimates of the RDR for blood pressure, total cholesterol and HDL cholesterol from previous studies. The blue circles correspond to the 4–, 16– and 20–year BRHS estimates presented in this chapter.



Chapter 6

Impact of within–person variation on CHD risk associations

6.1 Summary

In this chapter, the age–adjusted relations between the established coronary risk factors (blood lipids, blood pressure, body mass index, and lifestyle characteristics including cigarette smoking and physical activity) and major CHD risk (coronary death or non–fatal myocardial infarction) over the first twenty years of the BRHS are estimated. Estimates of the relative hazard for these factors are presented before and after correction for within–person variation, and are presented separately by each decade of follow–up. The combined effects of regression dilution bias in more than one risk factor on estimated multivariate relationships are also assessed. The results indicate strong positive (negative for HDL) linear associations between different blood lipid and blood pressure indices and major CHD risk, and a weaker positive relationship for body mass index. After correction for regression dilution bias, major CHD risk increased by 63% for each mmol/l increase in total cholesterol, and increased by 48% for every 20 mmHg increase in systolic blood pressure. Before correction for regression dilution bias these estimates were 39% and 25% respectively. Taking within–person variation in categorical risk factors into account improved the predictiveness of these exposures. On average, cigarette smokers had approximately a twofold risk of major CHD over never smokers; this risk increased with amount of cigarettes smoked each day. Men who were moderately active had nearly a 70% lower risk of CHD than inactive men; those who were "light" drinkers had a 22% lower risk than

non–drinkers. The estimated effects of these risk factors appeared to vary over the period of follow–up, though for the continuous risk factors, this dependency was largely abolished by performing "time–dependent" correction for regression dilution bias. Estimated regression coefficients in a multivariate analysis varied little depending on whether "multivariate" or "univariate" correction methods were employed, though some small differences were observed.

6.2 Introduction and objectives

In chapter 5, the extent of within–person variation in continuous and categorical risk factors over a twenty year period was assessed. For the established continuous risk factors, this was described in terms of estimating the regression dilution ratio (RDR) over a variety of follow–up periods and then approximating the nature of the relationship through these estimates by fitting a negative exponential curve (see Figures 5.2 and 5.3). For the categorical risk factors, the follow–up questionnaires (Q5, Q92, Q96 and Q20) were used to derive "average" exposures to these factors over the period of risk. For these factors, it was shown that baseline risk factor levels tended to be more extreme than the true underlying exposures to these factors. In this chapter, the effects of this variation on estimated relationships between established CHD risk factors and major CHD risk (coronary death or non–fatal myocardial infarction) over 20 years are examined.

For the continuous risk factors (in particular the blood lipid and blood pressure indices and body mass index), the estimated relationships between the RDR and the interval period over which it is estimated are used to relate 20–year major CHD risk with levels of these factors at the midpoint of this interval (i.e. after 10 years).[19] Relations between categorical risk factors and 20–year major CHD risk are estimated, both before and after adjustment for within–person variation. The relative ability of different blood lipid and blood pressure measures to predict CHD risk, as well as the relative "superiority" of using average lifestyle risk factors in analyses rather than baseline levels is also assessed. Due to the long period of follow–up of study participants, the degree to which risk factor associations are constant over time (the proportional hazards assumption) is also assessed by comparing risk associations during the first and second decades of follow–up. Finally, the impact of within–person variability in two risk factors on the adjusted CHD relations

of several factors (including factors measured precisely) are assessed. This is done by estimating the associations between total cholesterol and systolic blood pressure with CHD risk after adjustment for each other, age and cigarette smoking (current, ex, never) for men with no baseline diagnosis of CHD.

6.3 Methods

This chapter examines the relations between the established coronary risk factors and major CHD risk, defined as coronary death or non–fatal myocardial infarction, over 20 years of follow–up. Four different blood lipid indices (total, LDL, HDL, and the ratio of total to HDL) and four blood pressure indices (systolic, diastolic, mean arterial pressure, and mid blood pressure) are selected and compared in terms of strength of association and predictive ability. Pulse pressure (the difference between systolic and diastolic pressure) was not considered as evidence from other studies suggests that this measure is a relatively poor predictor of CHD risk (compared with the other blood pressure indices).[113] Lifestyle characteristics before and after adjustment for within–person variation (baseline measurements *vs* measurements that take into account follow–up questionnaires) are also compared in terms of strength of association and predictive ability. Each of these risk factors is considered first in isolation (adjusted only for age). In section 6.4.8, the effect of within–person variation in several risk factors on estimated multivariate associations is considered. In this analysis, only men with no baseline evidence of CHD are included.

6.3.1 Kaplan–Meier curves and relative "informativeness"

Kaplan–Meier curves stratified by baseline fifth of blood lipids (total, LDL, HDL and total:HDL), blood pressure (systolic, diastolic, mean arterial and mid blood pressure), and body mass index were used to display differences in the cumulative incidence of major CHD events over 20 years by levels of these variables; tests for difference were based on the log–rank test. Kaplan–Meier curves were also used to display differences in the cumulative incidence of CHD by cigarette smoking exposure, physical activity level and alcohol intake, both before and after taking within–person variation (follow–up questionnaires) into account. The "relative informativeness" of different blood lipid and blood pressure indices (divided into fifths) was assessed by comparing their respective log–rank statistics.

Similarly, the relative informativeness of baseline versus "average" exposures to cigarette smoking, physical activity and alcohol intake was ascertained by comparing the magnitude of the log–rank statistics before and after adjustment for within–person variation (i.e. comparing the log–rank statistic calculated using baseline exposures to that obtained when using the "average" exposures).

6.3.2 Relationship between baseline levels and 20–year major CHD risk

Cox proportional hazards regression was used to estimate the age adjusted relative hazard of major CHD over 20 years for each fifth of the baseline distributions of blood lipids, blood pressure and body mass index, and for baseline categories of cigarette smoking, physical activity and alcohol intake level. Relative hazards corresponding to unit increases in BMI, blood lipid and blood pressure levels were estimated by fitting these terms as continuous variables in an age–adjusted model. These are presented per 1 mmol/l increase in total and LDL cholesterol, per 20 mmHg increase in systolic blood pressure, per 10 mmHg increase in diastolic, mean arterial and mid blood pressure, and per 1 kg/m^2 increase in BMI. For HDL cholesterol and the ratio of total to HDL cholesterol, analyses were performed on the log scale and hazard ratios are presented per 20% increase in these factors (calculated by $1.2^{\hat{\beta}}$, where $\hat{\beta}$ is the estimated log hazard ratio). For presentation purposes, the proportional hazards regression coefficients relating log relative hazard to each these risk factors (by fifth of the blood lipid, blood pressure and BMI distributions and for each category of cigarette smoking, physical activity and alcohol intake) are presented as "floating absolute risks" [477] (see Figures 6.1 to 6.9).

6.3.3 Relationship between "usual" levels and 20–year major CHD risk

For the continuous measurements, 20–year major CHD risks were related to estimated levels at the midpoint of this interval (i.e. after ten years) by estimating the 10–year regression dilution ratios for these variables (see Table 6.1), calculating the expected 10–year level conditional on the baseline level for each individual (see equation 3.19), and recalculating the hazard ratios and floating absolute risks from these data. For the categorical risk factors, hazard ratios and floating absolute risks were calculated for the estimated "average" exposure categories rather than the baseline categories.

6.3.4 Proportional hazards assumption

For each of the CHD risk factors, separate estimates of the hazard ratio were calculated over the first and second ten years of follow–up, in order to assess the magnitude of possible deviations from the proportional hazards assumption by decade of follow–up. For the continuous risk factors (blood lipids, blood pressure and body mass index), this was examined:

1. By relating CHD risk in both decades to usual risk factor levels at the midpoint of the twenty year interval (1988/90);

2. By relating CHD risk in each decade separately to usual risk factor levels during that decade (time–dependent correction for regression dilution bias),[533] i.e. risk in the first decade was related to usual risk factor levels approximately 5 years after baseline and risk in the second decade was related to usual risk factor levels approximately 15 years after baseline.

Formal tests of non–proportionality were performed through assessment of the Schoenfeld residuals (as described in section 3.4.2). For cigarette smoking, physical activity and alcohol intake, it would have been problematic to estimate decade–specific risk associations in the same manner as for the continuous risk factors, not least because no information was available at ten years of follow–up. Rather, the bulk of the information about these risk exposures was collected after the first decade of follow–up had already elapsed.

6.3.5 Multivariate associations

In multivariate analyses where two or more variables are subject to within–person variation, estimated regression coefficients may under or overestimate true regression coefficients depending on the degree of correlation between the variables.[26;27] Rosner's multivariate correction method (described in chapter 3), estimates the vector of true regression coefficients $\boldsymbol{\beta}^*$, from the vector of observed regression coefficients $\boldsymbol{\beta}$, using the equation $\boldsymbol{\beta}^* = \boldsymbol{\beta}\boldsymbol{\Lambda}^{-1}$, where $\boldsymbol{\Lambda}$ is a matrix of regression coefficients relating follow–up levels of all measurements (including those measured precisely) to observed baseline levels. In chapter 5, estimates of $\boldsymbol{\Lambda}^{-1}$ were presented to describe the relations between a blood lipid measurement (total or HDL cholesterol), a blood pressure measurement (systolic or diastolic pressure), and two variables that can be thought of as precisely measured (age and

cigarette smoking status, classified as never, ex, or current), for men with no baseline evidence of CHD followed over 20 years. These estimates showed that the combination of using total cholesterol and systolic blood pressure yielded the least bias in estimation of other associations. In this chapter, this estimate of $\mathbf{\Lambda}^{-1}$ is used to estimate the true relative hazard of major CHD 15–25 years after baseline, for age, total cholesterol, systolic blood pressure, and cigarette smoking status, after multivariate correction for regression dilution bias in total cholesterol and systolic blood pressure. The results from four models are presented:

1. an analysis that ignores the effects of regression dilution bias entirely;
2. an analysis that corrects for regression dilution bias in total cholesterol only;
3. an analysis that corrects for regression dilution bias in systolic blood pressure only; and
4. an analysis that corrects for regression dilution bias in both total cholesterol and systolic blood pressure.

95% confidence intervals for the regression–dilution–corrected hazard ratios were obtained by using bootstrap resampling (with 1000 replicates) to estimate the 2.5 and 97.5 bias–corrected percentiles of the distribution of the corrected coefficients (see section 3.4.4).

6.4 Results

6.4.1 Population characteristics

Table 6.2 shows the baseline characteristics of the 7,735 men in the British Regional Heart Study together with the estimated usual risk factor levels over the 20–year study period. Mean total cholesterol was 6.3 mmol/l, mean LDL cholesterol 4.2 mmol/l, mean HDL cholesterol 1.12 mmol/l and the mean ratio of total to HDL cholesterol was 5.6. Mean blood pressure was 145/82 mmHg and mean BMI was 25.5 kg/m^2. After taking within–person variation in lifestyle risk factors into account (as described in chapter 5), 1,802 men (23.4%) were categorised as "never smokers" throughout the study, 2,525 men (32.8%) as "ex–smokers since 1978/80", and 207 men (2.7%) as "new" or "recurrent" cigarette smokers. Of the remaining men, the vast majority smoked an average of between 1 and 20

cigarettes a day (2,489 men), with 601 men (7.8%) smoking 21–39 cigarettes a day and just 95 men (1.2%) truly smoking at least 40 cigarettes a day. For physical activity, after taking within–person variation into account, the modal exposure categories were "occasional" and "light", with approximately 28% of men falling into each category. Relatively few men were either inactive throughout the study (6.8%) or vigorously active throughout the study (4.8%). For alcohol intake, nearly 70% of all men were either occasional or light drinkers throughout the study, with a further 19% of men classified as "moderate drinkers". Few men (3.8%) were truly heavy drinkers based on average alcohol intake during the study. Over the first twenty years of follow–up, 1,299 men (16.8%) experienced a major CHD event. The age–adjusted relations between the established coronary risk factors and major CHD risk are now described.

6.4.2 Blood lipids and major CHD

The age–adjusted relationships between blood lipids and 20–year CHD risk (both before and after correction for regression dilution bias) are shown in Figures 6.1 and 6.2. For serum total cholesterol and LDL cholesterol, the Kaplan–Meier CHD event curves and floating absolute risks (by fifth of the baseline distribution) display a clear positive continuous relationship with major CHD risk over 20 years. Before correction for regression dilution bias, the estimated relative hazard corresponding to a 1 mmol/l increase in total cholesterol was 1.39 (95% confidence interval 1.32 to 1.46). After correction for regression dilution bias, this hazard ratio increased to 1.63 (95% CI 1.52 to 1.76). For LDL cholesterol, the relative hazard of major CHD per 1 mmol/l increase was 1.41 (95% CI 1.32 to 1.50) before, and 1.74 (95% CI 1.57 to 1.93) after correction for regression dilution bias. HDL cholesterol displayed a continuous negative relationship with major CHD risk (Figure 6.2). For each 20% increase in HDL cholesterol, major CHD risk was estimated to decrease by 18% (95% CI 14% to 21%) before correction and by 21% (95% CI 16% to 25%) after correction for regression dilution bias. The ratio of total to HDL cholesterol displayed a similarly strong positive relationship with major CHD risk as was observed for total cholesterol. For each 20% increase in the ratio of total to HDL cholesterol, major CHD risk was estimated to increase by 29% (95% CI 25% to 34%) before correction and by 40% (95% CI 34% to 47%) after correction for regression dilution bias. For each of the blood lipid indices, there was no evidence of any "threshold" level below which a lower

level did not confer a lower (respectively higher for HDL) risk of major CHD, though it is acknowledged that these analyses would have had only limited power to detect such a threshold, if one did exist. The log–rank statistics by fifth of the distributions of the different blood lipid indices were highest for total cholesterol and the ratio of total to HDL cholesterol ($\chi^2 = 161$ and 174 respectively), and lowest for HDL cholesterol ($\chi^2 = 68$).

6.4.3 Blood pressure and major CHD

The age–adjusted relationships between blood pressure and 20–year CHD risk (before and after correction for regression dilution bias) are shown in Figures 6.3 and 6.4. Continuous graded positive relationships are seen for all blood pressure indices, with no evidence of any threshold levels below which a lower exposure does not confer a lower CHD risk. Before correction for regression dilution bias, the risk of major CHD was estimated to increase by 25% for each 20 mmHg increase in systolic pressure, by 19% for each 10 mmHg increase in diastolic pressure and mean arterial pressure, and by 17% for each 10 mmHg increase in mid blood pressure. After correction for regression dilution bias, a 20 mmHg increase in systolic pressure was estimated to increase CHD risk by 48% (95% CI **36%** to 61%), while a 10 mmHg increase in diastolic pressure and mean arterial pressure was estimated to increase CHD risk by 56% (95% CI 42% to 72%) and 42% (95% CI 33% to 53%) respectively. A 10 mmHg increase in mid blood pressure was associated with a 36% (95% 28% to 44%) increase in major CHD risk after correction for regression dilution bias. The relative informativeness of the different blood pressure indices was highest for systolic pressure (log-rank $\chi^2 = 149$) and lowest for diastolic pressure (log-rank $\chi^2 = 83$).

6.4.4 Body mass index and major CHD

The relationship between body mass index and major CHD risk is shown in Figure 6.5. The risk of major CHD decreased steadily with decreasing BMI, at least down to a level of 21 kg/m^2. However, the strength of relationship was weaker than that observed for the blood lipid or blood pressure indices, with less than a twofold difference in risk between those in the top and bottom fifths of the distribution. The effects of regression dilution bias for body mass index were minimal, due to the stability of this measure over time. For a 1 kg/m^2 increase in BMI, the risk of major CHD was estimated to increase by 6% (96% CI 4% to 8%).

6.4.5 Comparison of aetiological force between different risk factors

Figure 6.6 shows age–adjusted estimates of the hazard ratio for major CHD for each of the blood lipid and physical measures (blood pressure and body mass index), by comparing the risks of those in the top fifth of the distribution with the risks of those in the bottom fifth (vice versa for HDL cholesterol). The difference in major CHD risk across the risk factor distribution is greatest for the blood lipid indices (particularly total cholesterol and the ratio of total to HDL cholesterol) where approximately a threefold difference in risk is observed, and is lowest for body mass index (less than a twofold difference). The different blood pressure indices displayed almost identical gradients in major CHD risk across their distributions (approximately a twofold difference in major CHD). For the blood lipid indices, the gradient in major CHD risk was notably lower for HDL cholesterol than for the other blood lipid measures.

6.4.6 Lifestyle characteristics and major CHD

The age–adjusted relationships between cigarette smoking, physical activity and alcohol intake and major CHD risk over 20 years are shown in Figures 6.7 to 6.9, both before and after correction for within–person variation (changes in lifestyle recorded on the follow–up questionnaires).

Cigarette smoking

Figure 6.7 shows the relationship between cigarette smoking and major CHD risk. Relative to individuals who had never smoked cigarettes, ex–smokers had a 42% higher risk of CHD according to baseline data. After taking within–person variation into account, this decreased to a 37% excess risk, due to the prior inappropriate inclusion of a number of "new/recurrent" smokers amongst the "ex–smokers" (a group that had a 75% excess risk of CHD over never smokers). When baseline data were used to define smoking exposure, CHD risk varied little by amount smoked – the relative hazard of CHD compared with never smokers was 1.90 (95% CI 1.60 to 2.25) for those who smoked 1–20 cigarettes a day, 2.34 (95% CI 1.92 to 2.86) for those who smoked 21–39 cigarettes a day, and 1.79 (95% CI 1.32 to 2.42) for those who smoked 40 or more a day. However, when reported changes in cigarette smoking habits over time were taken into account, a clear dose response relationship between CHD risk and amount smoked was observed. Compared with never

smokers, the true relative hazard of CHD was 1.71 (95% CI 1.45 to 2.02) for those who smoked 1–20 cigarettes a day, 3.05 (95% CI 2.48 to 3.76) for those who smoked 21–39 cigarettes a day, and 4.17 (95% CI 2.81 to 6.20) for those who smoked 40 or more a day. The informativeness of baseline cigarette smoking exposure was substantially lower than the informativeness of a smoking exposure category that took follow–up questionnaires into account (log–rank statistic equalled 105 before and 149 after correction for within–person variation).

Physical activity

Figure 6.8 shows the relationship between physical activity and major CHD risk. Risk was highest amongst inactive men, and decreased progressively with increasing levels of physical activity up to moderate levels, after which a small increase in risk was observed. Relative to inactive men, the age–adjusted relative hazard of major CHD for moderately active men was 0.47 (95% CI 0.42 to 0.65) before correction and 0.32 (95% CI 0.25 to 0.39) after correction for within–person variation. The relative risk reduction corresponding to vigorous levels of physical activity was similar before and after correction for within–person variation, the vigorously active had approximately half the risk of major CHD of the inactive. The informativeness of the six–level physical activity scale was vastly improved by taking within–person variation into account (the log–rank test statistic increased from 82 to 194).

Alcohol intake

The relationship between alcohol intake (defined as none, occasional, light, moderate or heavy) and major CHD risk is displayed in Figure 6.9. A U-shaped relationship was observed by baseline levels of alcohol intake, with men who were non–drinkers at baseline having the highest subsequent rates of major CHD. "Light" drinkers had the lowest observed risks of major CHD, 34% (95% CI 18% to 47%) lower than than those of non-drinkers. However, after taking account of within–person variation in alcohol intake, the risks of non–drinkers were not significantly different from those of either occasional or moderate drinkers, and the estimated benefits from "light" levels of drinking were reduced from 34% to 22% (95% CI 6% to 36%). Furthermore, men who were truly heavy drinkers throughout the study had a 75% higher risk of CHD (95% CI 31% to 133%)

over non–drinkers (compared with an estimated 16% lower risk from baseline data). As was observed for cigarette smoking and physical activity, the relative informativeness of the five–level alcohol intake scale was increased after taking within–person variation into account (the log–rank test statistic increased from 22 to 56).

6.4.7 Non–proportionality and time–dependent correction for regression dilution bias

Table 6.3 shows the relationships between usual risk factor levels over the twenty–year period and the risk of major CHD in each decade of follow–up (1978/80 – 1988/90 and 1988/90 to 1998/00). It can be seen that for the blood lipid and blood pressure indices, the strength of relationship between levels ten years after baseline (1988/90) and major CHD risk is stronger for the first ten years than the second. The apparent "non–proportionality" of these indices is confirmed by a formal test of the Schoenfeld residuals (see last column of Table 6.3). Similarly, for cigarette smoking and alcohol exposure, some degree of non–proportionality is observed, though relative hazard estimates corresponding to physical activity were fairly stable over the follow–up period. However, by relating CHD risk in each decade to usual exposure levels after ten years of follow–up, no account is taken of the fact that the interval between the baseline measurement and the date of first major CHD event is substantially shorter for individuals who have their event in the first decade than for individuals who have their event in the second. The average "time to event" for the 652 men who experienced a first major CHD event between 1978/80 and 1988/90 was 5.3 years. In comparison, for the 647 men who experienced their first major CHD event between 1988/90 and 1998/00, it was 14.9 years. Therefore, in order to properly assess whether blood lipid and blood pressure effects are constant over the period of follow–up, one should relate CHD risk in the first decade to usual risk factor levels approximately 5 years after baseline, and relate CHD risk in the second decade to usual risk factor levels approximately 15 years after baseline. The effect of performing such "time–dependent" correction for regression dilution bias is shown in Table 6.4, where it can be seen that the hazard ratio estimates are now fairly similar between the two decades, indicating that the effects of usual blood lipids, blood pressure and BMI on CHD risk are actually reasonably constant throughout the study.

6.4.8 Multivariate adjustment

The effects of adjustment for regression dilution in two major imprecisely measured CHD risk factors, systolic blood pressure and serum total cholesterol, their relations with CHD outcome – and with the influence of other factors generally regarded as precisely measured – age, and smoking status, were examined in 6,576 men initially free from CHD. Relative hazards for major coronary heart disease risk in the long term (15–25 years after baseline) adjusted for age, systolic blood pressure, total cholesterol and smoking status are shown in Table 6.5 both before and after correcting for the effects of regression dilution at 20 years. Four models are presented, all of which assume that age and smoking status are known precisely. The first model corresponds to an analysis using baseline measures alone, and ignores the effects of regression dilution for both total cholesterol and systolic blood pressure. This analysis provides the 'naïve' hazard ratio estimates corresponding to a 20 mmHg increase in SBP and a 1 mmol/L increase in cholesterol of 1.31 and 1.35 respectively. The second model corrects for the effects of regression dilution for SBP only. This results in an increase of the adjusted hazard ratio for SBP to 1.76, with little change to the other hazard ratios. Similarly, the third model corrects solely for regression dilution of total cholesterol, increasing its hazard ratio to 1.75. Again, relatively little change in the hazard ratios for the other factors is observed. Correction for regression dilution of both total cholesterol and systolic blood pressure is shown in the final model, where hazard ratio estimates for SBP and total cholesterol increase to 1.94 and 1.77 respectively. For systolic pressure, this hazard ratio is greater than that obtained after correction for regression dilution in SBP alone. In model 4, where multiple correction for regression dilution in both SBP and total cholesterol is performed, the hazard ratios for age and cigarette smoking status change very little.

6.5 Discussion

6.5.1 Interpretation of findings

Over twenty years of follow–up, strong continuous relations were observed between blood lipid and blood pressure indices and major CHD risk; the magnitudes of these associations were increased after correction for regression dilution bias. Major CHD risk increased by 63% for each mmol/l increase in usual total cholesterol and by 48% for every 20 mmHg

increase in usual systolic blood pressure (respectively 56% per 10 mmHg increase in usual diastolic pressure). Body mass index also displayed a consistent, albeit weaker, positive relationship with CHD risk, with risk increasing by 6% for every kg/m^2 increase in BMI. The ability to separate high and low–risk individuals based on levels of cigarette smoking, physical activity and alcohol intake was greatly improved when information from follow–up questionnaires was taken into account. In particular, the large excess risks of heavy smoking and the full benefits of moderate physical activity only became truly apparent after within–person variation in these factors was adjusted for. The use of average levels of alcohol exposure in analyses (rather than baseline levels) reduced the size of the apparent benefits on CHD risk from light levels of drinking, and identified the potential risks from regular heavy drinking which were not previously observed. The associations between average risk exposure levels and major CHD risk appeared to vary over the period of follow–up, though for the continuous risk factors, this variation was largely abolished after time–dependent correction for regression dilution bias was performed.

When examining the multivariate effects of age, total cholesterol, systolic blood pressure and cigarette smoking on CHD risk in men with no prior evidence of CHD, correction for regression dilution bias in both systolic blood pressure and total cholesterol had little effect on the hazard ratio estimates for age or cigarette smoking. However, correction for regression dilution in both factors led to a slightly higher hazard ratio for systolic pressure than was observed after correction for regression dilution bias in systolic pressure alone.

6.5.2 Validity of methods

For the continuous risk factors, correction for regression dilution bias was performed by estimating the regression dilution ratio for each factor over a ten–year interval, so that CHD risk over the period 1978/80 to 1998/00 could be related to risk factor levels at the midpoint of this interval (ten years after baseline). This approach to correction for regression dilution bias has previously been employed in several studies of cardiovascular disease.[17;19;113;533–535] One aspect in which the analyses presented in this chapter differ from these previous studies however, is that the 10–year regression dilution ratios used were not estimated from real data over a 10 year period. Rather, they were predicted from the estimated functional relationship between the RDR (estimated over periods of one week, and 4–, 16– and 20 years) and the interval period (see Table 6.1). While it would

have been preferable to base these adjustments on estimates of the RDR taken from real data over a ten year period, the consistency of the relationships between the RDR and the interval (see Chapter 5; Figures 5.2 and 5.3), as well as the agreement between the 10–year estimates derived from this prediction approach and the ten–year estimates obtained from real data in other studies (see Chapter 5; Figure 5.12), suggest that the estimates used in this chapter are likely to be valid. In order to examine the influence of these risk factors by decade of follow–up, this correction approach was subsequently extended to allow for "time–dependent" correction for regression dilution bias. This was performed by estimating the 5– and 15–year RDR estimates so that risk in the first decade could be related to five–year risk factor levels and risk in the second decade could be related to 15–year risk factor levels, an approach that has previously been used in other large prospective studies.[113;533]

For the categorical risk factors, an "averaging approach" (described in section 5.3.3 and illustrated in Figure 5.1) was used to define exposure categories that take into account information supplied by the follow–up questionnaires. The key feature of this approach that greatly affects its validity in this setting is that only information on risk factor changes that was obtained while the individuals were *still at risk of a first major CHD event* were used to derive the summary average exposures. This is important in order to eliminate the potential for reverse causality bias, i.e. the onset of disease causing a change in subsequent risk exposures. Though men with pre–existing CHD were also included in the analyses (a group likely to have already had their risk exposures modified because of CHD), the relationships between their risk exposures and their risk of further major CHD remain relevant to the assessment of CHD risk factor associations in the population of all middle–aged British men, and hence they were kept in the analyses. In any case, the impact of keeping these men in analyses was small; very similar results were obtained when they were excluded from analyses. Similar approaches of taking into account within–person variation in lifestyle coronary risk factors (particularly physical activity) have been used elsewhere.[190;536]

6.5.3 Comparison with other studies

The associations between the established CHD risk factors and the occurrence of major CHD presented in this chapter are generally consistent with those of previous studies,

though many of these are based on fatal CHD rather than all major CHD, and most (though not all) have not corrected for within–person variation (regression dilution bias). A brief comparison between the findings in this chapter and those of other large studies (and meta–analyses of studies) is now provided. A detailed description of the epidemiological evidence regarding each risk factor is given in chapter 2.

Blood lipids, blood pressure and body mass index

In the Prospective Studies Collaboration, blood pressure was related to the occurrence of fatal CHD in nearly one million adults from 61 separate cohort studies (including the BRHS), and time–dependent correction for regression dilution bias was performed.[113] In men aged 60-69, a 20 mmHg increase in systolic blood pressure was found to be associated with an 82% increase in the risk of CHD death. This estimate is somewhat higher that the estimate obtained in this chapter (48%), which may, at least partially, be due to the different endpoint used in the PSC analysis. For total cholesterol, a study published in 1994 consisting of ten prospective studies, three international studies, and 28 randomised controlled trials estimated that at age 60 (the average age of the men in the BRHS after 10 years of follow–up), a reduction in total cholesterol of 1 mmol/l was associated with a 41% reduction in the risk of fatal CHD.[70] This estimate (which was corrected for regression dilution bias in total cholesterol) is almost exactly the same as the estimate presented in this chapter (where a 1 mmol/l decrease in total cholesterol was associated with a 39% reduction in major CHD risk; since $1/1.63 = 0.61$). Our estimates of the relationship between body mass index and CHD risk are also in general agreement with previous publications. For instance, in Finnish men aged 30–59 years, a unit increase in BMI was associated with a 4% increase in the risk of CHD[226] (compared with a 6% increase presented in this chapter). A recent meta–analysis of the effects of body mass index on CHD risk in Chinese adults also estimated a 7% increase in risk per 1 mg/m^2 increase in BMI.[229]

Cigarette smoking, physical activity and alcohol

In a case–control study of 14,000 survivors of myocardial infarction (and 32,000 controls), cigarette smokers had about five times the risk of myocardial infarction than non–smokers at age 40–49, three times the risk at age 50–59 and two and a half times the risk at age

60–69.[151] In the BRHS, when all cigarette smokers were considered together, a twofold difference in risk of major CHD was observed between current smokers and never–smokers. In the 40–year report from the British Doctors Study,[145] the rate of fatal CHD was 56% higher in cigarette smokers than non–smokers (892 *vs* 572 deaths per 100,000 men per year). For physical activity, a meta–analysis of the effects of physical activity on CHD risk published in 1990 found that those in "active" occupations had approximately half the risk of CHD of those in sedentary occupations.[173] These data have recently been supported by a review of the benefits of physical activity on the risk of CHD, which concluded that being physically active was associated with about a 40 to 50% reduction in the risk of CHD.[196] However, most of these previous studies used baseline measures in analyses. The conclusion that risk can be halved through moderate levels of physical activity is consistent with the estimate in this chapter obtained from baseline measurements (see Figure 6.8). However, after taking within–person variation in physical activity into account, moderately active men were observed to have approximately a 70% lower risk of CHD than inactive men. These estimates of the potential true benefits of physical activity are difficult to assess in the context of other studies however, due to the few studies that have attempted to control for within–person variation. However, one study that used repeated measurements of physical activity exposure to quantify "physical activity group" found that women who (on average) walked for at least two hours a week had a 67% lower risk of CHD than women who did not walk regularly.[182] Men whose usual physical activity level was "moderately vigorous" or "vigorous" appeared to have a greater risk of major CHD than men who exercised moderately (see Figure 6.8), a phenomena that has also been observed in the Harvard College Alumni Study[172] and the Multiple Risk Factor Intervention Trial.[194] The mechanism behind this increased risk remains unclear, though (in the BRHS at least) it does not appear to be mediated through an increased risk of sudden cardiac death.[196]

As observed for physical activity, the age–adjusted relationship between baseline alcohol intake and CHD risk was consistent with previous studies that have used baseline measures in analyses.[255;256] Baseline alcohol intake displayed a U–shaped relationship with CHD risk, with men who drank light to moderate amounts of alcohol experiencing around 25–30% lower CHD risks than men who did not drink. However, after correction for within–person variation, a smaller protective effect of moderate levels of alcohol was observed and a previously unobserved excess risk associated with heavy drinking was

identified. These findings are difficult to assess in the context of other studies however, as few authors have commented on the likely effects that within–person variation in alcohol consumption could have on estimated alcohol–CHD relationships. [i]

6.5.4 Conclusions: 20–year associations with major CHD

In the British Regional Heart Study, continuous linear relationships were observed between blood lipid and blood pressure indices and major CHD risk over 20 years. After correction for regression dilution bias, a 1 mmol/l increase in serum total cholesterol was associated with 63% increase in CHD risk and a 20 mmHg increase in systolic pressure was associated with a 48% increase in CHD risk. Lifestyle risk factors, including cigarette smoking, physical activity and alcohol intake, were improved as explanatory factors when information from follow–up questionnaires was taken into account. After correction for within–person variation, clear dose–response relationships were observed between amount of tobacco smoked and CHD risk, and the estimated benefits of regular physical activity during middle–age increased. For alcohol intake however, the apparent benefits from light to moderate levels of drinking were attenuated when within–person variation was taken into account, and the group of regular heavy drinkers were shown to have a substantially higher CHD rate than the non–drinkers. In a multivariate setting, within–person variation in more than one factor had some residual effects on the estimated importance of other factors, including those measured precisely, though these effects were relatively small.

[i] While the potential impact of within–person variation on alcohol–CHD relationships was discussed in a report from the British Doctors' Study,[260] the authors were unable to directly assess these effects, though a reasonable degree of consistency between alcohol intake at the beginning and end of the study (for surviving men) suggested that their results may have been fairly robust to these effects.

Table 6.1: Predicted regression dilution ratios for blood lipids, blood pressure and body mass index over a ten year interval

Risk factor	Relationship between RDR and interval period (years)	RDR(10)
Blood lipids		
Serum total cholesterol	$\mathrm{RDR}(t) = \exp(-0.175 - 0.023t)$	0.67
LDL cholesterol	$\mathrm{RDR}(t) = \exp(-0.228 - 0.026t)$	0.61
HDL cholesterol*	$\mathrm{RDR}(t) = \exp(-0.036 - 0.014t)$	0.84
Ratio (total to HDL cholesterol)*	$\mathrm{RDR}(t) = \exp(-0.090 - 0.018t)$	0.76
Blood pressure		
Systolic blood pressure	$\mathrm{RDR}(t) = \exp(-0.383 - 0.016t)$	0.58
Diastolic blood pressure	$\mathrm{RDR}(t) = \exp(-0.411 - 0.051t)$	0.40
Mean arterial pressure†	$\mathrm{RDR}(t) = \exp(-0.384 - 0.033t)$	0.49
Mid blood pressure‡	$\mathrm{RDR}(t) = \exp(-0.380 - 0.027t)$	0.52
Body mass index	$\mathrm{RDR}(t) = \exp(0.014 - 0.004t)$	0.97

* Analysed on the log scale
†(2/3)DBP + (1/3)SBP
‡(SBP + DBP)/2

Table 6.2: Risk factor levels recorded for all men at baseline and estimated true risk factor levels over the period of follow–up. Data indicate mean (SD) or n (%) unless otherwise stated

Risk exposure	Baseline level	True level ‡
Blood lipids		
Total cholesterol (mmol/L)	6.3 (1.0)	6.3 (0.9)
LDL cholesterol (mmol/L)	4.2 (1.0)	4.2 (0.8)
HDL cholesterol (mmol/L)†	1.12 (0.96 - 1.29)	1.12 (0.98 - 1.26)
Total to HDL ratio†	5.6 (4.6 - 6.7)	5.6 (4.8 - 6.4)
Blood pressure		
Systolic pressure (mmHg)	145 (21)	145 (16)
Diastolic pressure (mmHg)	82 (13)	82 (8)
Mean arterial pressure (mmHg)	103 (15)	103 (10)
Mid blood pressure (mmHg)	114 (16)	114 (11)
Body mass index (kg/m^2)	25.5 (3.2)	25.5 (3.2)
Cigarette smoking		
Never smoked cigarettes	1819 (23.6%)	1802 (23.4%)
Ex-cigarette smoker	2715 (35.2%)	2525 (32.8%)
New cigarette smokers	NA	207 (2.7%)
Current smoker (1-20 a day)	2023 (26.2%)	2489 (32.3%)
Current smoker (21-39 a day)	846 (11.0%)	601 (7.8%)
Current smoker (40 or more a day)	316 (4.1%)	95 (1.2%)
Physical activity		
None	686 (9.0%)	517 (6.8%)
Occasional	2345 (30.7%)	2138 (28.0%)
Light	1761 (23.1%)	2152 (28.2%)
Moderate	1205 (15.8%)	1426 (18.7%)
Moderately vigorous	1120 (14.7%)	1030 (13.5%)
Vigorous	513 (6.7%)	367 (4.8%)
Alcohol		
None	466 (6.0%)	755 (9.8%)
Occasional	1845 (23.9%)	2101 (27.2%)
Light	2544 (32.9%)	3141 (40.6%)
Moderate	2042 (26.4%)	1441 (18.6%)
Heavy	832 (10.8%)	291 (3.8%)

†Geometric mean (interquartile range);
‡Estimated level after taking within–person variation into account;

Table 6.3: Estimated age adjusted relative hazard of major CHD over the first and second ten years of follow–up after adjustment for within–person variation.*

Risk factor	First ten years (652 events)		Second ten years (647 events)		p†
	HR	95% CI	HR	95% CI	
Blood lipids					
Total cholesterol (1 mmol/L)	1.70	(1.53,1.88)	1.57	(1.42,1.74)	0.02
LDL cholesterol (1 mmol/L)	1.85	(1.60,2.13)	1.64	(1.41,1.90)	0.02
HDL cholesterol (20% increase)	0.76	(0.71,0.82)	0.84	(0.77,0.91)	0.01
Total to HDL ratio (20% increase)	1.46	(1.37,1.55)	1.33	(1.24,1.43)	<0.01
Blood pressure					
Systolic pressure (20 mmHg)	1.56	(1.39,1.75)	1.39	(1.23,1.57)	0.08
Diastolic pressure (10 mmHg)	1.70	(1.49,1.95)	1.42	(1.24,1.64)	0.04
Mean arterial pressure (10 mmHg)	1.51	(1.37,1.67)	1.33	(1.20,1.48)	0.04
Mid blood pressure (10 mmHg)	1.43	(1.31,1.55)	1.29	(1.18,1.41)	0.05
Body mass index (kg/m^2)	1.07	(1.04,1.09)	1.06	(1.03,1.09)	0.22
Cigarette smoking					0.05
Never smoked cigarettes	1.00	–	1.00	–	
Ex-cigarette smoker	1.48	(1.15,1.90)	1.28	(1.01,1.61)	
New cigarette smokers	2.03	(1.25,3.30)	1.53	(0.93,2.52)	
Current smoker (1-20 a day)	1.74	(1.35,2.22)	1.70	(1.35,2.13)	
Current smoker (21-39 a day)	3.39	(2.53,4.54)	2.71	(2.00,3.67)	
Current smoker (40 or more a day)	3.78	(2.15,6.64)	4.70	(2.70,8.21)	
Physical activity					0.85
None	1.00	–	1.00	–	
Occasional	0.58	(0.45,0.73)	0.64	(0.49,0.85)	
Light	0.41	(0.32,0.53)	0.41	(0.31,0.55)	
Moderate	0.29	(0.21,0.39)	0.35	(0.26,0.49)	
Moderately vigorous	0.35	(0.26,0.49)	0.42	(0.30,0.58)	
Vigorous	0.47	(0.31,0.71)	0.54	(0.35,0.82)	
Alcohol					<0.01
None	1.00	–	1.00	–	
Occasional	1.26	(0.94,1.69)	0.93	(0.72,1.19)	
Light	0.99	(0.74,1.32)	0.64	(0.49,0.82)	
Moderate	1.51	(1.11,2.04)	0.79	(0.60,1.05)	
Heavy	2.49	(1.68,3.68)	1.22	(0.79,1.87)	

* for continuous risk factors, major CHD related to levels after 10 years of follow–up.
†test of the Schoenfeld residuals for non–proportionality.

Table 6.4: Estimated age adjusted relative hazards of major CHD over the first and second ten years of follow–up after time–dependent correction for regression dilution bias †

Risk factor	First ten years (652 events)		Second ten years (647 events)	
	HR	95% CI	HR	95% CI
Blood lipids				
Total cholesterol (1 mmol/L)	1.60	(1.46,1.75)	1.66	(1.48,1.87)
LDL cholesterol (1 mmol/L)	1.71	(1.51,1.94)	1.75	(1.48,2.08)
HDL cholesterol (20% increase)	0.77	(0.72,0.83)	0.83	(0.76,0.90)
Total to HDL ratio (20% increase)	1.41	(1.33,1.50)	1.37	(1.27,1.48)
Blood pressure				
Systolic pressure (20 mmHg)	1.51	(1.36,1.68)	1.42	(1.25,1.63)
Diastolic pressure (10 mmHg)	1.51	(1.36,1.68)	1.58	(1.32,1.90)
Mean arterial pressure (10 mmHg)	1.42	(1.31,1.54)	1.40	(1.24,1.59)
Mid blood pressure (10 mmHg)	1.37	(1.27,1.47)	1.33	(1.20,1.48)
Body mass index (kg/m^2)	1.07	(1.04,1.09)	1.06	(1.03,1.09)

†Major CHD risk related to usual risk factor levels 5 years after baseline for the first decade and 15 years after baseline for the second decade.

Table 6.5: Estimated relative hazard of major CHD after 15–25 years of follow–up before and after correcting for the effects of regression dilution in serum total cholesterol and systolic blood pressure

	(a) HR	95% CI	(b) HR	95% CI	(c) HR	95% CI	(d) HR	95% CI
Age (years)	1.07	(1.05,1.08)	1.05	(1.04,1.06)	1.07	(1.06,1.09)	1.06	(1.04,1.07)
Smoking status								
Non-smoker	1.00	-	1.00	-	1.00	-	1.00	-
Ex-smoker	1.16	(0.96,1.42)	1.18	(0.95,1.45)	1.17	(0.95,1.42)	1.19	(0.94,1.48)
Current smoker	1.97	(1.64,2.36)	2.01	(1.68,2.46)	2.02	(1.69,2.50)	2.06	(1.71,2.57)
Usual SBP (20 mmHg)	1.31	(1.23,1.38)	1.76	(1.55,2.01)	1.36	(1.28,1.45)	1.94	(1.67,2.24)
Usual TC (1 mmol/L)	1.35	(1.28,1.43)	1.36	(1.28,1.45)	1.75	(1.57,1.96)	1.77	(1.59,2.01)

SBP = systolic blood pressure;
TC = serum total cholesterol;
HR = hazard ratio;
CI = confidence interval;
(a) Uncorrected for regression dilution;
(b) Corrected for regression dilution in SBP only;
(c) Corrected for regression dilution in TC only;
(d) Corrected for regression dilution in SBP and TC;

Figure 6.1: Age–adjusted associations between total and LDL cholesterol and 20–year major CHD risk before (black circles) and after (white circles) correction for regression dilution bias. Hazard ratio estimates correspond to unit (1 mmol/l) increases in cholesterol level.

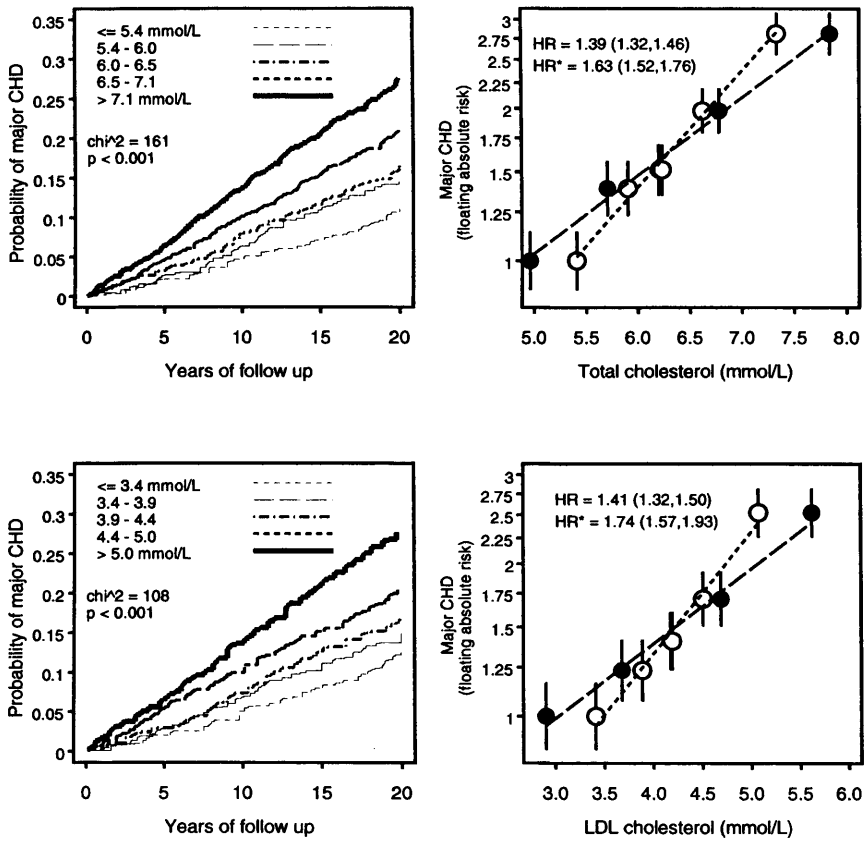


Figure 6.2: Age–adjusted associations between HDL and the ratio of total to HDL cholesterol and 20–year major CHD risk before (black circles) and after (white circles) correction for regression dilution bias. Hazard ratio estimates correspond to 20% increases in cholesterol level.

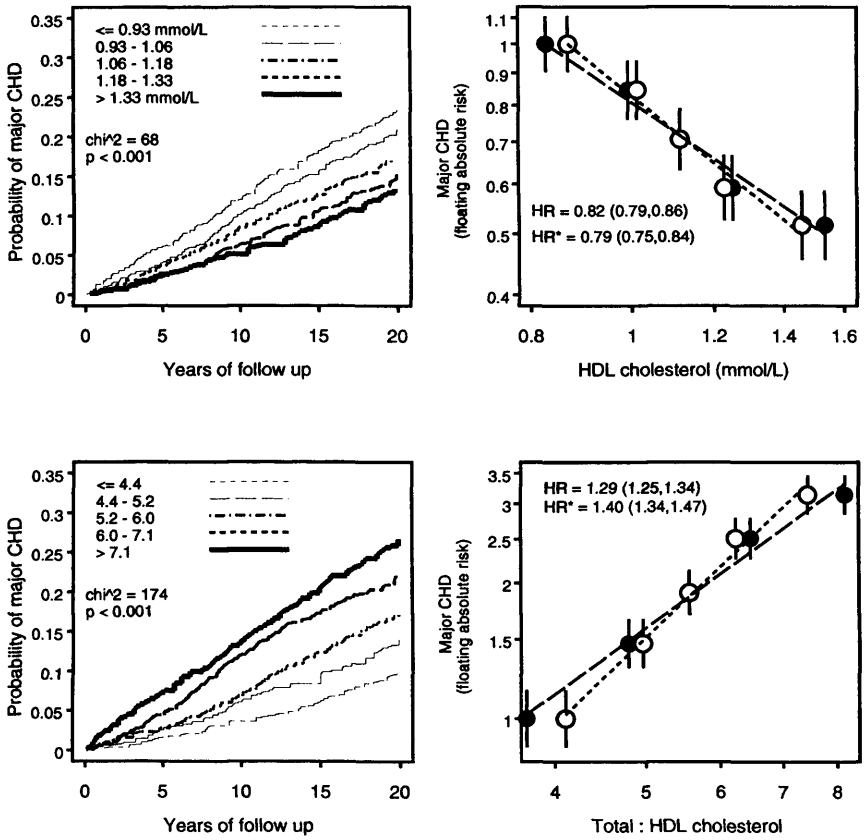


Figure 6.3: Age–adjusted associations between systolic and diastolic blood pressure and 20–year major CHD risk before (black circles) and after (white circles) correction for regression dilution bias. Hazard ratio estimates correspond to increases of 20 mmHg for systolic, and 10 mmHg for diastolic pressure.

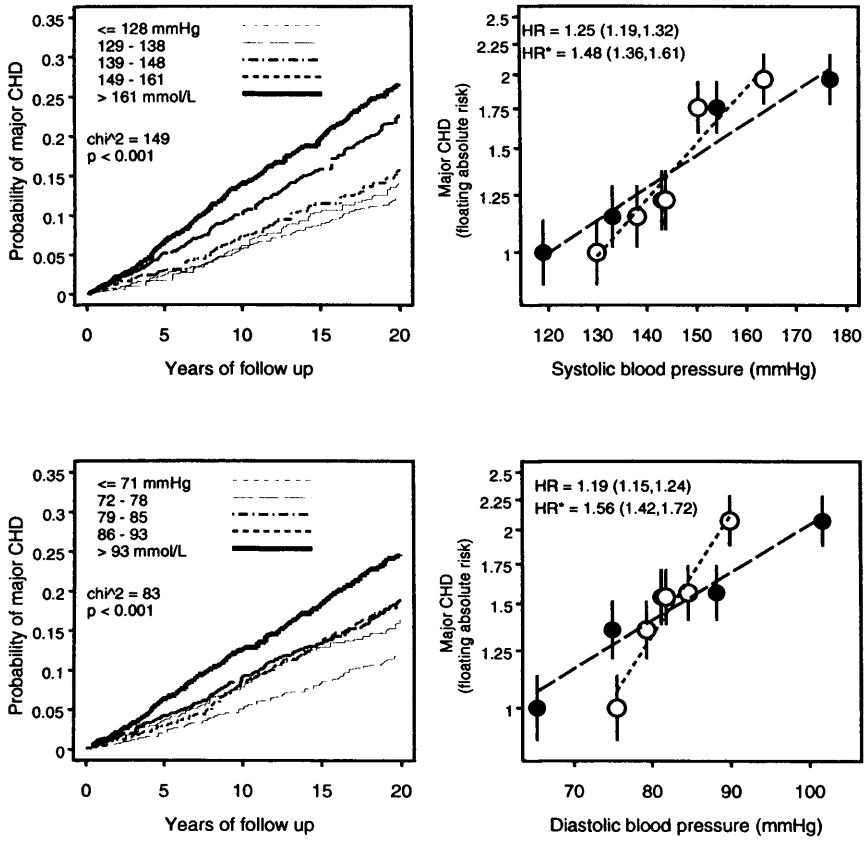


Figure 6.4: Age–adjusted associations between mean arterial and mid blood pressure and 20–year major CHD risk before (black circles) and after (white circles) correction for regression dilution bias. Hazard ratio estimates correspond to 10 mmHg increases in pressure.

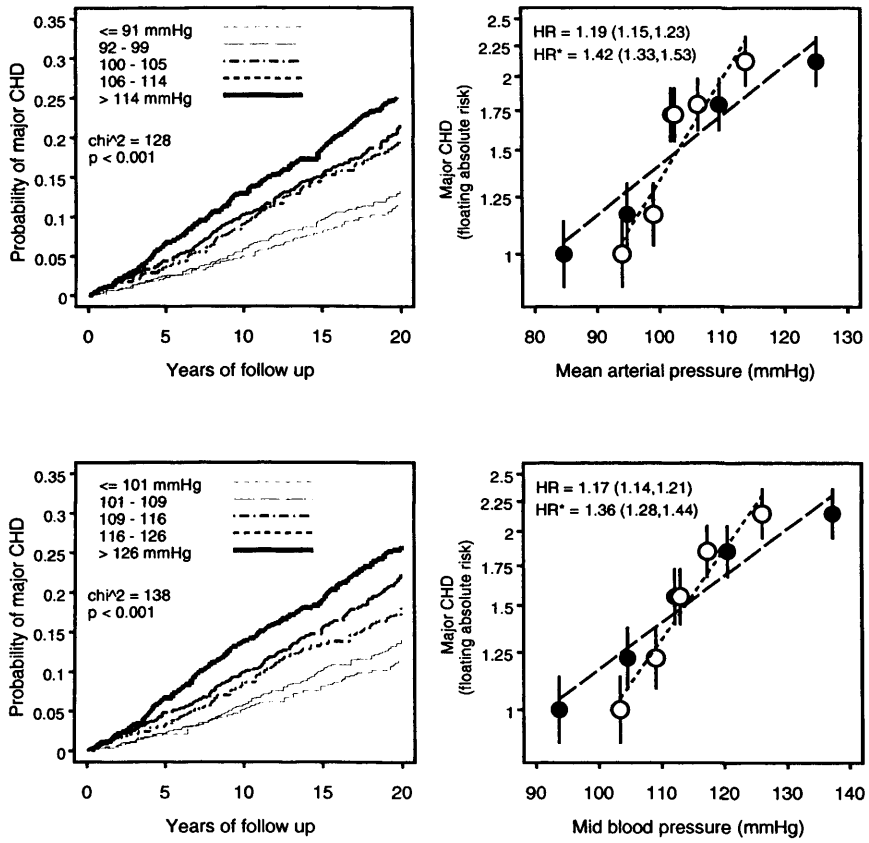


Figure 6.5: Body mass index and 20–year major CHD risk before (black circles) and after (white circles) correction for regression dilution bias.

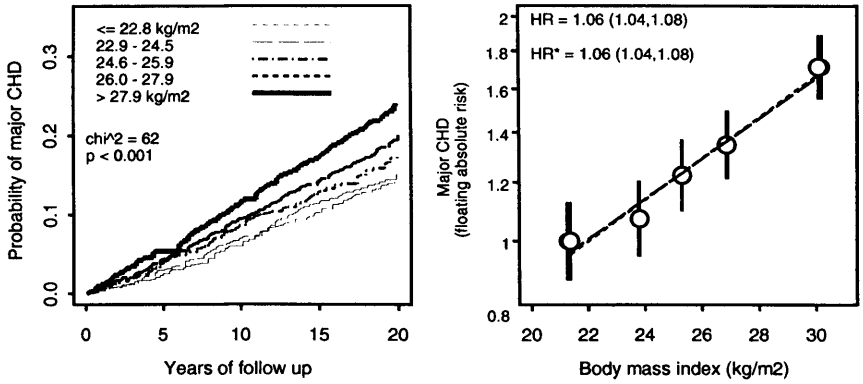


Figure 6.6: Age adjusted relative hazard of major CHD within 20 years. Comparisons shown are between the top and bottom fifths of the distributions of each factor (for HDL cholesterol this is the bottom fifth relative to the top fifth).

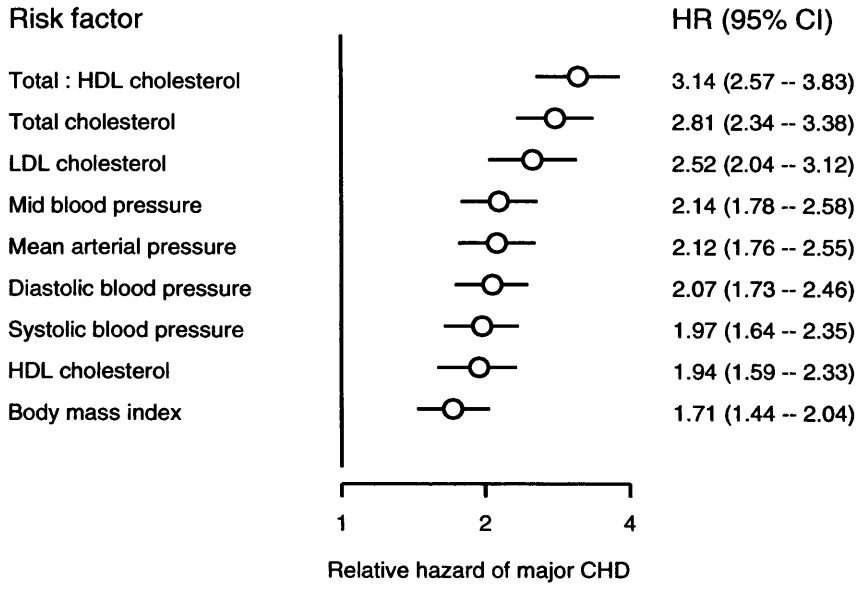
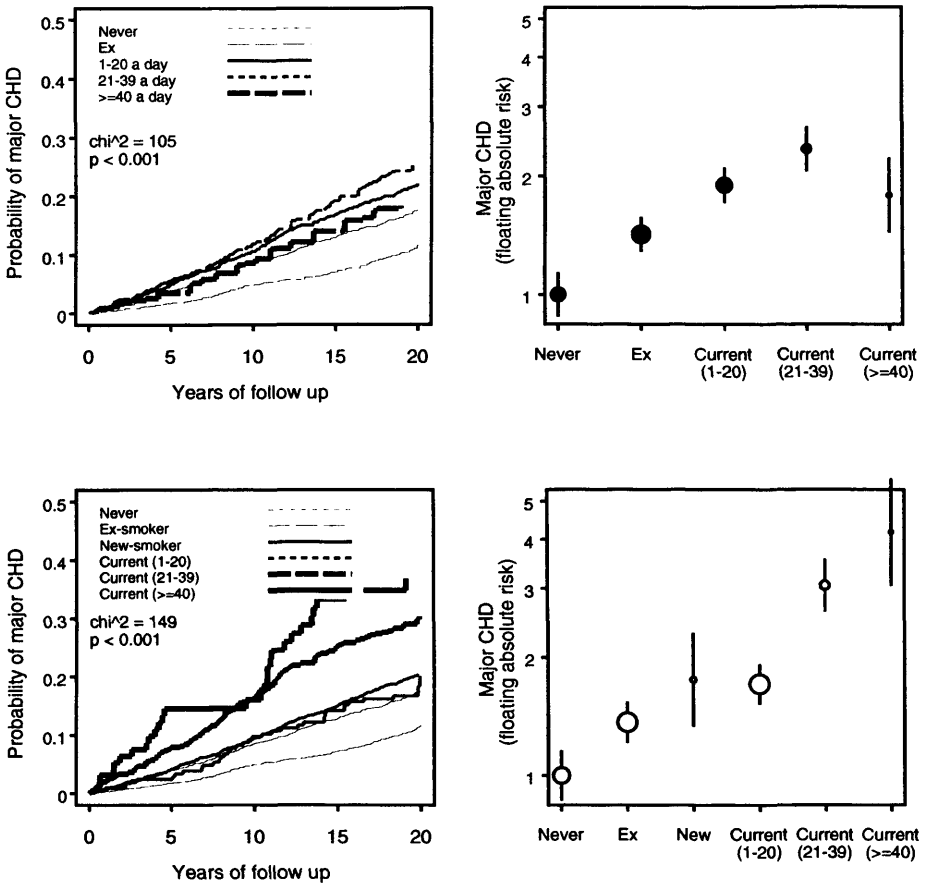


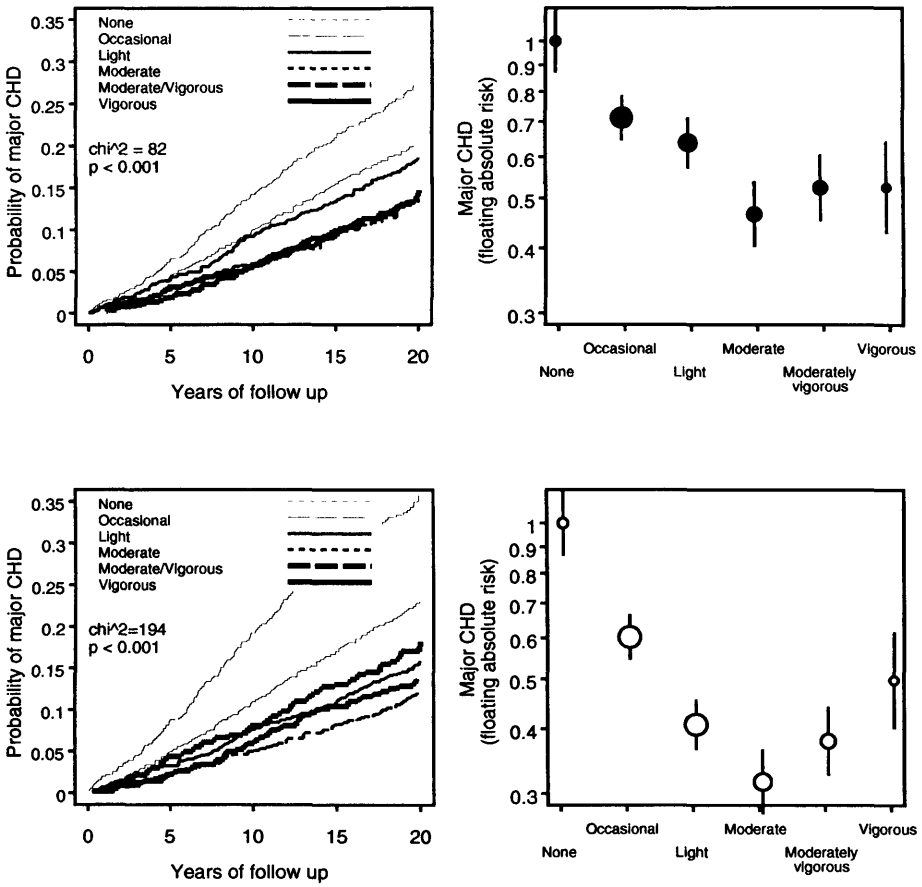
Figure 6.7: Cigarette smoking and 20–year major CHD risk before (top) and after (bottom) taking into account information from follow–up questionnaires.



Cigarette smoking group	HR[1] (95% CI)	HR[2] (95% CI)
Never smoked cigarettes	1.00	1.00
Ex cigarette smoker	1.42 (1.20,1.68)	1.37 (1.15,1.62)
New cigarette smoker	NA	1.75 (1.24,2.48)
Current smoker (1–20 a day)	1.90 (1.60,2.25)	1.71 (1.45,2.02)
Current smoker (21–39 a day)	2.34 (1.92,2.86)	3.05 (2.48,3.76)
Current smoker (40 or more a day)	1.79 (1.32,2.42)	4.17 (2.81,6.20)

1. HR estimates from baseline data
2. HR estimates taking into account within–person variation

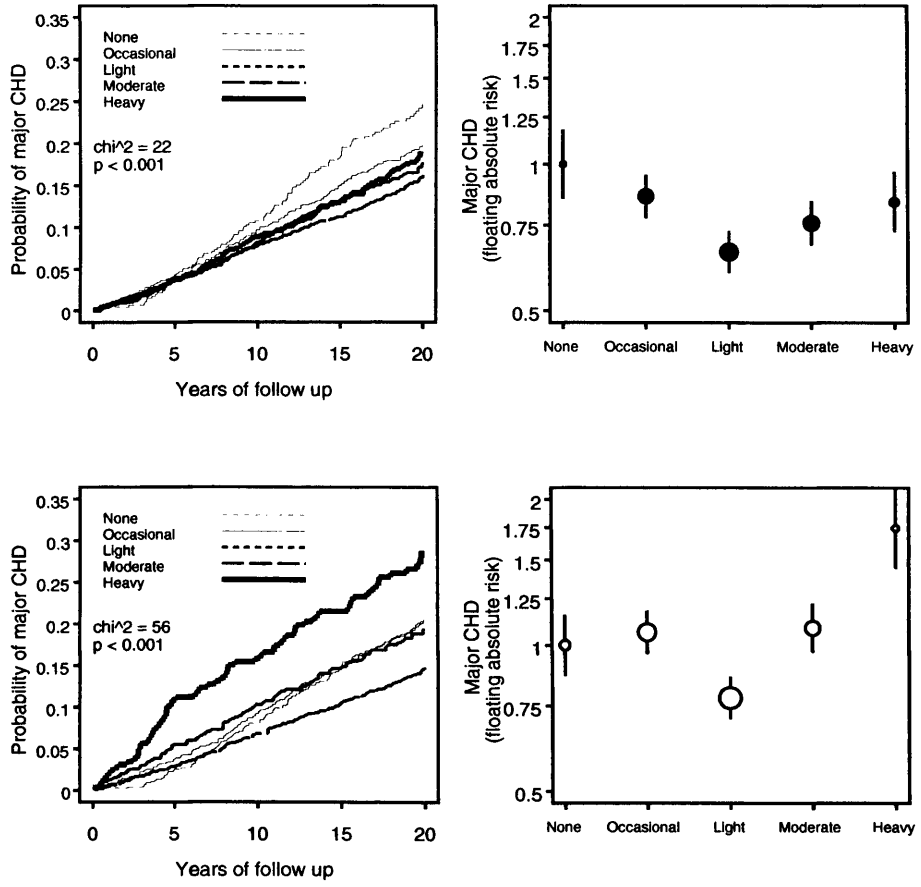
Figure 6.8: Physical activity and 20–year major CHD risk before (top) and after (bottom) taking into account information from follow–up questionnaires.



Physical activity group	HR[1] (95% CI)	HR[2] (95% CI)
None	1.00	1.00
Occasional	0.71 (0.60,0.85)	0.60 (0.50,0.72)
Light	0.64 (0.53,0.77)	0.41 (0.34,0.49)
Moderate	0.47 (0.37,0.58)	0.32 (0.25,0.39)
Moderately vigorous	0.52 (0.42,0.65)	0.38 (0.30,0.48)
Vigorous	0.52 (0.40,0.69)	0.50 (0.37,0.67)

1. HR estimates from baseline data
2. HR estimates taking into account within–person variation

Figure 6.9: Alcohol and 20–year major CHD risk before (top) and after (bottom) taking into account information from follow–up questionnaires.



Alcohol intake	HR[1] (95% CI)	HR[2] (95% CI)
None	1.00	1.00
Occasional	0.86 (0.69,1.07)	1.06 (0.88,1.29)
Light	0.66 (0.53,0.82)	0.78 (0.64,0.94)
Moderate	0.76 (0.61,0.95)	1.09 (0.88,1.33)
Heavy	0.84 (0.65,1.08)	1.75 (1.31,2.33)

1. HR estimates from baseline data
2. HR estimates taking into account within–person variation

Chapter 7

Re–assessing the "only 50%" claim

7.1 Summary

The aim of this chapter is to assess the validity of the claim that only half of all CHD cases may be "attributed" to its three strongest causal risk factors: blood cholesterol, blood pressure and cigarette smoking. The relationships between these factors and major CHD events are examined over a ten–year period for the 6,576 men with no baseline evidence of CHD. Analysis is carried out before and after multivariate correction for regression dilution bias in blood cholesterol and blood pressure, and the estimated risk associations are used to assess the potential combined contribution of these factors to population attributable CHD risk. In men with no pre–existing evidence of CHD, clear positive relationships between usual blood cholesterol and blood pressure levels and major CHD risk were observed, with no evidence of any threshold below which a lower level did not confer a lower risk. When the CHD risk of non–smoking individuals with "low" cholesterol and "low" blood pressure was compared with those of all remaining "high–risk" individuals, large differences were observed. In direct contrast to the "only 50% claim", the combined population attributable risk fraction for "high" cholesterol, "high" blood pressure and cigarette smoking ranged from around 70% to over 90% (depending on how the low–risk (reference) group was defined) after correction for regression dilution bias. Narrowing the low–risk group to exclude ex–smokers and passive smokers increased relative risk and population attributable risk estimates still further. Adjustment for other independent coronary risk factors, including body mass index, physical activity and social class had little effect on the results.

7.2 Introduction

7.2.1 Background

The claim that the three strongest risk factors for coronary heart disease (blood cholesterol, blood pressure and cigarette smoking) account for only approximately half of CHD risk (the "only 50% claim"), has become widely accepted as fact,[30–40] though it is often stated with no supporting data,[30;32;34;38] or asserted with inappropriate citations.[31;33;35] Furthermore, despite evidence from large prospective studies suggesting that this figure may be an underestimate,[41;43–47] the claim is still being made,[35–39] and is even repeated in reviews presented as "state of the art".[38] The belief that half of CHD risk is explained by the established risk factors is crucially important. Not only does it suggest that the scope for prevention of CHD by altering exposure to these established risk factors is limited (reducing it by half, at most), but also that there must be an unexplained 50% of CHD risk, thus implying the existence of important undiscovered risk factors.[42]

7.2.2 Objectives

The objective in this chapter is to assess the validity of the "only 50% claim" for men in the British Regional Heart Study, before and after correction for within–person variation in CHD risk factors. In particular, the combined population attributable risk fraction (PARF) for "high" blood cholesterol, "high" blood pressure and current cigarette smoking is estimated, taking account of regression dilution bias in both blood pressure and blood cholesterol. Since the relations between blood cholesterol and blood pressure and CHD risk are continuous and "threshold free", separation of individuals into "high" and "low" risk groups based on these risk factors is arbitrary. Therefore, estimates of the PARF are presented as curves (or functions) of the cut–off criteria used to define the high–risk group. In a subsidiary analysis the effects of using different thresholds for cigarette smoking exposure taking account of previous active smoking and passive smoking, and the effects of adjustment for other CHD risk factors are examined.

7.3 Subjects and methods

The analyses in this chapter are restricted to the 6,576 men with no baseline evidence or symptoms of CHD (defined as recall of a doctor diagnosis of myocardial infarction or angina, Rose chest pain questionnaire evidence of definite or possible angina, or a history of severe chest pain; see Chapter 4, Table 4.5). This is to reduce the likelihood of associations being influenced by reverse causality (i.e. pre–existing disease causing a change in the risk factor). In addition, in contrast to the previous chapters, only major CHD events over the first ten years of follow–up (rather than for 20 years) are included. This is for a number of reasons:

1. By using ten years rather than twenty years of follow–up, the four–year repeated measurements in Dewsbury and Maidstone (between 1996 and 2000) can be used to perform multivariate correction for regression dilution bias in both the blood lipid and the blood pressure measurement, thus allowing 10–year risk to be related to risk exposure levels at (approximately) the midpoint of the interval. Note that while the individual regression dilution ratios could also have been predicted over a ten–year interval (as was done in Chapter 6 when univariate adjustments were performed), this would not allow assessment of the potential effects following multivariate correction for regression dilution bias.

2. When examining associations over 20 years (in Chapter 6), some "non–proportionality" was observed for each of the risk factors considered in this chapter. Though this was overcome (for the continuous risk factors) by performing "time–dependent" correction for regression dilution bias, such an approach could not easily be adopted in the current multivariate setting without additional repeated risk factor measurements.

3. For consistency with the following chapter (that examines different approaches to the primary prevention of CHD) it is advantageous in this chapter to:

   (a) reduce the effect that preventive drugs would have on observed CHD rates by concentrating on events between 1978/80 and 1988/90 (these drugs would predominantly (exclusively for statins) have been taken during the 1990s, i.e. the second decade of follow–up),

(b) concentrate exclusively on CHD events occuring during middle–age (as these events may be considered as preventable rather than (possibly) inevitable), and

(c) use a risk exposure–period over which current primary prevention policies recommend evaluating absolute CHD risk (i.e. ten years).

7.3.1 Association between usual risk factor levels and major CHD risk

The associations between baseline risk factors and ten–year major CHD risk were estimated through Cox proportional hazards regression; analyses were adjusted for age, blood cholesterol, blood pressure and cigarette smoking status (current *vs* ex/never). It was assumed throughout that blood cholesterol and blood pressure were measured with error and subject to variation over time, but that cigarette smoking was known precisely (though some within–person variation in cigarette smoking exposure was identified in chapter 5, this was predominantly observed between different levels of current smoking, rather than between the current *vs* ever/never exposure groups considered in this chapter). The effects of within–person variation in blood lipids and blood pressure on 10–year disease relationships were assessed using the four–year repeated data (from Dewsbury and Maidstone in 1996 and 2000). As described in chapter 5 (and illustrated in chapter 6), regression coefficients may underestimate or overestimate true regression coefficients when two or more variables are measured with error.[521] The multivariate techniques developed by Rosner *et al.*[26] (described in section 3.7) were therefore used to correct the 10–year baseline associations for (multivariate) regression dilution bias. This method estimates the true regression coefficients $\boldsymbol{\beta}^*$ from the baseline coefficients $\boldsymbol{\beta}$ through the equation $\boldsymbol{\beta}^* = \boldsymbol{\beta}\boldsymbol{\Lambda}^{-1}$, where $\boldsymbol{\Lambda}^{-1}$ is the "modifying matrix" described in chapter 5. 95% confidence intervals for $\boldsymbol{\beta}^*$ were calculated using bias–corrected bootstrap re–sampling of size 1000.

7.3.2 Choosing the "best" blood cholesterol and blood pressure indices

Some indices of blood cholesterol and blood pressure may be better or worse at predicting major CHD events than others. In the age–adjusted analyses presented in Chapter 6, for instance, the Kaplan–Meier log–rank statistics (by fifth of the baseline distributions) provided one means of assessing the "informativeness" of a risk factor. In this chapter, the relative abilities of different blood cholesterol and blood pressure indices to predict CHD risk is assessed through examination of the overall contribution that different baseline

measurements make to the χ^2 likelihood ratio statistic in the Cox proportional hazards regression model.[113] Three blood lipid indices (total cholesterol, HDL cholesterol and the ratio of total to HDL) and three blood pressure indices (systolic, diastolic and mean arterial pressure) were selected and compared in this way. LDL cholesterol was not considered in this analysis because it was only measured in 18 of the 24 towns studied. The results for mid blood pressure were virtually identical to those of mean arterial pressure, and are therefore not presented separately in this chapter.

7.3.3 Separation of individuals into risk groups

To determine population attributable risk fractions it is necessary to distinguish a "low risk" group (assumed to be free of the relevant risk exposures) from one or more "high risk" groups in which the exposures are present. However, the relationships between blood cholesterol, blood pressure and CHD are continuous and threshold free,[73;113] and so division of these factors into risk categories is arbitrary. In the primary analysis, a range of potential "cutoff" values have therefore been used, including the usual 10th and 20th centiles of the risk exposure distribution in the population, and the target levels currently recommended in the United Kingdom for preventive treatment with lipid lowering and blood pressure lowering drugs (total cholesterol 5 mmol/L and blood pressure 140/85).[408] Relative risk and PARF–curves are subsequently presented to describe the nature of the relationship between these statistics and the cut–off levels used to define the high–risk group. For cigarette smoking, "current smokers" have been used as the high–risk group in the main analysis; the effects of extending this group to include "ex–smokers" and subjects exposed to passive smoking are considered in subsidiary analyses.

7.3.4 Predicting the relative risk (RR) and the PARF

Cox proportional hazards regression coefficients (before and after multivariate correction for regression dilution bias) were used to predict the 10–year risk of major CHD for any given level of total cholesterol, blood pressure and smoking status. For a particular cut–off criterion, the expected 10–year risk of major CHD for "high–risk" men was then calculated by taking the average over all possible high–risk men (obtained by numerically integrating the predicted risk distribution over the theoretical range of values defining the high–risk group). The expected 10–year risk of major CHD for "low–risk" men was calculated

similarly, allowing calculation of the expected relative risk of major CHD for high–risk men relative to low–risk men. This was done both before and after taking regression dilution bias into account. The population attributable risk fraction corresponding to the high–risk factors was then calculated by estimating, before and after correction for regression dilution bias, the proportion of the population at high risk, and using Levin's equation (chapter 3, equation 3.1) to calculate the PARF. Approximate 95% prediction intervals were calculated by calculating a large ($B = 1000$) bootstrap sample $\{\boldsymbol{\beta}_i^B\}$ of $\boldsymbol{\beta}$ and $\{\boldsymbol{\Lambda}_i^B\}$ of $\boldsymbol{\Lambda}$, repeating the above procedure for each sample, and taking the 2.5 and 97.5 bias–corrected percentiles as estimates of the lower and upper limits (see chapter 3, section 3.4.3).

Identifying the contribution of separate high–risk groups

The method used to calculate the population attributable risk fraction for a single high risk group above can easily be used to calculate the marginal PARFs for each of the different possible high–risk groups (see Chapter 3, equations 3.2 and 3.3). In the analyses presented in this chapter where an individual can be defined to be "high–risk" based on any one of three conditions being satisfied (high cholesterol, high blood pressure or cigarette smoker), there are seven possible high risk groups that could be considered separately (e.g. high cholesterol only, high blood pressure only, high blood pressure and smoker, etc.). These are listed in Table 7.5. Labeling these groups with the indices 1 to 7, the marginal PARF of the jth high risk group is estimated by

$$\text{PARF}_j = \frac{\hat{p}_j(\hat{R}R_j - 1)}{1 + \sum_{k=1}^{7} \hat{p}_k(\hat{R}R_k - 1)} \tag{7.1}$$

The sum of the seven marginal PARFs equals the overall PARF previously estimated.

7.4 Results

7.4.1 Usual risk factor levels during the first ten years

Of the 7,735 men examined at the baseline screening, 6,576 (85.0%) had no evidence of CHD at baseline (doctor diagnosis or symptoms of angina or myocardial infarction). Of the 361 men with no previous history of CHD who attended the 16–year screening in

Dewsbury and Maidstone (80% response), 259 (76% of survivors) also attended the 20–year r–screening. Estimates of the regression dilution ratio over this period (shown in Table 7.1) were calculated for each of the blood lipid and blood pressure indices. Using these values, the estimated usual risk factor levels in the population during the first decade of follow–up were then calculated from the observed baseline levels; these are shown in Table 7.2. Mean blood cholesterol and blood pressure levels were similar in smokers and non–smokers, and virtually no correlation between total cholesterol and blood pressure was observed ($\rho = 0.08$ for systolic, $\rho = 0.11$ for diastolic). Both serum total cholesterol and blood pressure (systolic, diastolic and mean arterial) were approximately normally distributed (see Figure 7.1); HDL cholesterol and the ratio of total to HDL cholesterol were normally distributed when plotted on the log scale.

7.4.2 Relative informativeness

The relative informativeness of the different blood cholesterol and blood pressure indices at predicting 10–year CHD risk was assessed in the subgroup of men with complete data on total and HDL cholesterol, blood pressure and a self reported cigarette smoking status (n =6,299, 96% of disease free population). This is shown in Table 7.3 where "informativeness" is expressed as a reduction in deviance from the model that adjusts only for age and smoking status and is presented as a percentage relative to the model that includes total cholesterol and systolic blood pressure. Relative to this reference model, most of the alternative models were inferior in terms of predictive ability. Mean arterial pressure was equal in terms of predictive ability to systolic pressure, and the logarithm of the ratio of total to HDL cholesterol was, if anything, slightly more predictive than total cholesterol alone. However, in the interests of having to take just one measure of each index (and, for the logarithm of total:HDL cholesterol, in the interest of a simple interpretation of the risk association), systolic pressure and total cholesterol were selected for use in the analyses that follow.

7.4.3 Relationships with 10–year major CHD

Of the 6,576 men with no baseline symptoms or diagnosis of CHD, 6,515 (99.1%) had complete baseline data on total cholesterol, systolic blood pressure and cigarette smoking status. Of these men, 426 (6.5%) had a definite major CHD event within the following 10

years. The relations between usual levels of serum total cholesterol and systolic blood pressure and CHD (adjusted for age, cigarette smoking, and each other) were approximately linear (Figure 7.2), and were unaffected by adjustment for body mass index, usual physical activity level, usual alcohol intake and history of diabetes. No significant non–linear or interaction effects between the variables were observed. After multivariate correction for regression dilution bias in total cholesterol and systolic pressure, major CHD risk was estimated to increase by 60% for a 1 mmol/l increase in total cholesterol and by 79% for a 20 mmHg increase in systolic pressure. Multivariate correction for regression dilution bias in both total cholesterol and systolic pressure caused a small reduction in the estimated hazard ratios for current cigarette smoking and age.

7.4.4 Relative and attributable risks – individual risk factors

Table 7.4 shows predictions of the relative risk and population attributable risk fraction associated with high serum total cholesterol and high systolic blood pressure (before and after correction for regression dilution bias) for the cases where the low risk group is defined by the 20th or 10th centiles of the usual risk factor distributions as well as when these are defined by thresholds currently used for clinical interventions. The estimates of the RR and the PARF as a function of the threshold criteria used to define the low–risk group is shown in Figure 7.3; each "RR–curve" and "PARF–curve" is adjusted for age and cigarette smoking and mutually adjusted for total cholesterol and systolic blood pressure. It can be seen that for threshold criteria set below the mean level in the population, the relative risk and PARF increase as the threshold criteria used to define the low–risk group decreases. The relative risks associated with high serum total cholesterol at the 20th and 10th centiles are 1.85 and 1.97 respectively, increasing to 2.10 and 2.31 after correction for regression dilution bias. At these levels, the PARFs for "high" serum total cholesterol are 39% and 45% before correction, and 47% and 54% after correction for regression dilution bias. If all men had experienced the risks of men with a total cholesterol of less than 4.5 mmol/L, then up to two–thirds of the major CHD events may have been avoided, however this should be stated cautiously given the very small proportion of men falling into the low–risk group at this level (approximately 3%). For systolic blood pressure, the relative risks of major CHD at the 20th and 10th centiles are 1.70 and 1.78, increasing to 2.30 and 2.58 after correction for regression dilution bias. This leads to PARF estimates of 35%

and 40% before, and 51% and 59% after correction. At the level above which intervention is recommended to reduce CHD risk, the PARF for high serum total cholesterol alone (≥5 mmol/L) is 48% (relative risk 2.02), increasing to 57% (relative risk 2.42) after correction for regression dilution, while for high systolic blood pressure (≥140 mmHg), the PARF (respectively RR) is 28% (1.64) before correction, and 41% (2.10) after. For current cigarette smoking, the PARF is 23% before correction for regression dilution bias. After correction for regression dilution bias in cholesterol and blood pressure, this figure was reduced to 21%.

7.4.5 Relative and attributable risk – combined risk factors

Figure 7.4 displays the RR and the PARF for high serum total cholesterol, high systolic blood pressure and current cigarette smoking considered simultaneously, both before and after correcting for the effects of regression dilution bias in total cholesterol and systolic pressure; estimates corresponding to specific criteria are shown in Table 7.4. Using the 20th centiles of the usual risk factor levels to define the low–risk group, and after correction for regression dilution bias, high risk individuals had an estimated risk of major CHD 5.24 times that of low–risk individuals and the PARF was 80%. Using the 10th centiles of usual risk factor levels to define the low risk group, the relative risk was 7.06 after correction for regression dilution and the PARF was 86%. For the intervention thresholds, the relative risk was 5.33 and the PARF 81% after correction for regression dilution bias.

Separating the high risk groups

Table 7.5 shows estimates of the marginal relative risks and PARFs for each of the seven separate high risk groups for two scenarios: (1) high–risk individuals are defined by the lowest tenths of the population distributions; and (2) high–risk individuals are defined by the lowest fifths. It can be seen that at these "cutoff" values, the majority of the population fall into the "high" total cholesterol and "high" blood pressure categories (with or without cigarette smoking), which leads to large marginal PARF estimates for these groups. In contrast, for high–risk groups that capture only a small proportion of the population (e.g. cigarette smoking and high blood pressure only), their contribution to overall population attributable risk is small despite large relative risks. For instance, after correction for regression dilution bias, a cigarette smoker with an SBP of at least 131 mmHg has, on

average, four times the risk of major CHD of a non–smoker whose SBP is 131 mmHg or lower, but the contribution of this group to all major CHD is small (around 3%), because only around 6% of the population fall into this group.

7.4.6 Subsidiary analyses

Varying the definition of cigarette smoking exposure

In the subset of 5,885 subjects who had either ever regularly smoked cigarettes or (if they had never been a smoker) had had their cotinine level recorded, the effect of using different low–risk thresholds for cigarette smoking on PARF estimates was examined (Figure 7.5). Restricting the low–risk group to subjects who had never smoked (i.e. adding ex–smokers to the high–risk group) had little effect on the relative risk or PARF estimates. However, restricting the low–risk group further to subjects who had never smoked and were not heavily exposed to environmental tobacco smoke (cotinine levels <1.5 ng/ml, excluding subjects exposed to a partner smoking 20 cigarettes or more per day)[537] increased estimates of the relative risk substantially (approximately a ten–fold difference in risk between the high and low–risk groups was predicted) and consequently the PARF estimates increased to approximately 90%. Use of the "usual" cigarette smoking exposure category (described in chapter 5) rather than the "baseline" exposure did not affect the results because all current cigarette smokers were considered as a single high–risk group in these analyses.

Effects of other CHD factors

Further analyses adjusting for other coronary risk factors (body mass index, usual physical activity level, usual alcohol intake, and history of diabetes), markers of deprivation in early and adult life (height and social class) and town of residence had very little effect on the relative risk and PARF estimates obtained, reducing them by approximately 3% (see Table 7.6). In addition, further exclusion of individuals with baseline ECG evidence of myocardial ischaemia or infarction (definite or possible) had no effect on the estimated combined population attributable risk fractions.

7.5 Discussion

7.5.1 Interpretation of findings

Using the 4–year repeated risk factor data, it has been possible to estimate the associations between usual risk factor levels over the first decade and major CHD risk over this period, after multivariate correction for regression dilution bias in both cholesterol and blood pressure. From this information, the relative risks and population attributable risk fractions associated with "high" total cholesterol, "high" blood pressure and cigarette smoking, both individually and simultaneously, have been estimated. These analyses allow estimation of the true contribution of blood cholesterol, blood pressure and cigarette smoking to CHD risk, providing a means of assessing the validity of the "only 50%" claim in the BRHS. Defining low–risk individuals as being in the lowest quintile (fifth) of usual levels of serum total cholesterol and systolic blood pressure and current non–smokers, the PARF was 70% before and 80% after correction for regression dilution bias. Therefore, had all individuals experienced the average risk of those in the low–risk group, four fifths of major CHD events within the following 10 years would have been avoided. Had everyone experienced the average risk levels of those in the bottom tenths of these distributions, 86% would have been avoided. Changing the definition of non–smokers to exclude ex–smokers and those exposed to heavy passive smoking increased these estimates still further to around 90%.

7.5.2 Validity of analyses

The validity of the analyses presented in this chapter depends on the appropriateness of the estimates of regression dilution used and the levels used to define individuals as "low risk". The estimates of regression dilution bias were based on repeat blood pressure and cholesterol measurements taken four years apart, and were restricted to individuals with no previous history of CHD. Though the subjects were older at the time of the repeat measurements than at baseline, estimates of regression dilution do not appear to vary markedly with age (as demonstrated in Chapter 5; Figures 5.6 to 5.11). Moreover, the 4–year regression dilution ratios of 0.72 and 0.62 for serum total cholesterol and systolic blood pressure respectively are consistent with those derived from other studies over similar periods (see Figure 5.12). By presenting relative risks and population attributable risk

fractions as "curves" or functions of the criteria used to calculate them, their relationships with the definition of the low–risk group have been identified. The separation of low and high risk groups on the basis of blood cholesterol and blood pressure levels is somewhat arbitrary, as evidence from other study populations strongly suggest that the relations between these factors and CHD risk have no threshold and continue below the levels of the bottom tenths of our study population.[71;73;113] Thus, even in the groups defined as "low–risk" in this chapter, the effects of cholesterol and blood pressure exposure on CHD risk are likely to be appreciable, compared for example with the Japanese cohorts of the Seven Countries Study, in whom both mean serum total cholesterol level and CHD mortality risk were markedly lower than was the case even in the lowest risk group considered here.[29] For example, an ecological study carried out in 1989 of diet, mortality and lifestyle in 69 nationally representative counties in rural China observed that average total cholesterol was 3.8 mmol/l, substantially lower than the "usual" average of 5.4 mmol/l estimated from men in the bottom fifth of the BRHS distribution, and that the death rate from CHD was only about one sixth of that in the UK. Reducing risk factors to levels below those experienced by the low–risk groups defined in this chapter (if possible) would therefore be likely to increase further the proportion of all CHD events prevented.

An alternative estimation approach

An alternative method of estimating the PARF corresponding to high levels of a continuous risk exposure would be to simply partition the data into two groups at the cut–off point and estimate the PARF directly from the data (either using classical methods or through fitting a binary term in a generalised linear model). However, in order to obtain the relationship between the PARF and the cut–off value used to define it (the "PARF–curve"), one would need to apply this technique at several cut–off levels and then fit a smooth function through the values. In addition, while methods exist for correcting binary risk factors for measurement imprecision,[21;22;24;493;538] (see chapter 3; section 3.6.2) these are usually based on the assumption that misclassification is non–differential. For a continuous symmetrically distributed exposure which is categorised as either "high" or "low" based on a cut–off level, this assumption is only likely to be true when the cut–off level is set to be equal to the mean. For all other cut–off levels (for instance levels below the mean), this assumption would be invalidated.

7.5.3 Comparison with other studies

The estimates of the PARF of 70–75% presented in this chapter (before correction for regression dilution bias) are consistent with estimates from earlier reports using the same risk factors. In the Multiple Risk Factor Intervention Trial (MRFIT), CHD death rates in over 270,000 men aged 40–57 years, who were initially free from disease, were compared between those at "low" risk (defined as serum total cholesterol < 5.17 mmol/L, blood pressure no higher than 120/80, and no cigarette smoking), and all remaining individuals. After 16 years of follow-up, the CHD death rate observed in the low–risk group was 78% lower than in the rest of the sample (PARF=77%).[44] The authors also estimated these risk differences for 7,490 middle–aged men and 6,229 middle–aged women who were free from CHD at initial screening into the Chicago Heart Association Detection Project in Industry study. After 22 years, CHD death rates in the low–risk group were 77% lower in men (PARF=76%), and 79% lower in women (PARF=78%), than CHD death rates in the high–risk groups. Even larger relative risks reductions were observed for men and women who were aged <40 years at time of recruitment. In each of these studies, low risk individuals comprised only approximately 5 to 10 percent of the cohorts, leading to the virtual equivalence between the relative risk reductions and the corresponding PARF values. Similarly, in the National Co–operative Pooling Project, risk of a first major coronary event was estimated to be lower by 70% for middle–aged men in the lowest quintile of risk (implying a PARF of approximately 65%), compared with all other men, [406] while recent results from the Framingham study estimate CHD risk to be considerably lower in low–risk men and women compared with all men and women.[407] In the United Kingdom, the Whitehall I study of over 17,000 middle–aged British civil servants found that if the average CHD mortality rate in the whole population could have been reduced to that experienced by individuals who had never smoked cigarettes and who were in the lowest quintiles of blood cholesterol and blood pressure levels, then about two–thirds of the CHD deaths would have been avoided.[42] None of these studies however, took into account the effects of regression dilution bias in blood cholesterol and blood pressure when making their estimates, though in some cases,[42;44] these effects were discussed. Most recently, the validity of the "50% claim" has been revisited from a different viewpoint. Two cardiovascular studies estimated the prevalence of established risk factors in a large number of individuals with coronary heart disease. The results showed that between 80

and 90% of such individuals were exposed to at least one of the main risk factors (high blood cholesterol, high blood pressure, smoking or diabetes).[46;47] While these findings also suggest that the "50% claim" may be somewhat of an underestimate, they are not as rigorous as the previously described studies (or indeed the analyses of this chapter) because of the possible influence that reverse causality bias may have played.

7.5.4 Within–person variation in cigarette smoking exposure

Though the analyses in this chapter have corrected for the effects of within–person variation in total cholesterol and blood pressure, they have not accounted for the possibility that the third factor, cigarette smoking exposure, is also subject to these influences. In chapter 5, it was observed that baseline cigarette smoking exposure was fairly representative of true "average" exposure to cigarette smoking throughout the study (with the possible exception of amount of cigarettes truly smoked). Therefore, it seems reasonable to assume that the categorisation of smoking exposure used in the analyses in this chapter (where all smokers are considered together) is appropriate. However, these analyses do not allow for the fact that by quitting smoking during the first ten years, some smokers may have influenced their CHD risk over that period (though it is more likely to take a few years for these benefits to be realised). In the BRHS, of the current smokers at baseline who survived event free for the following 5 years, 24% reported to have given up during that period. While this would not lead to misclassification of smoking status if current or ex–smoker were used as the defining high–risk group, if only current smoking were used (as in the primary analysis of this chapter), there is a potential for underestimation of the importance of current smoking.[78;168] This is because the risks of these individuals after quitting are lower than they would have been had they continued to smoke, and so by analysing them as if they remained smokers, true relative risks (and therefore population attributable risks) associated with current cigarette smoking may have been underestimated (though as already stated, this would depend on how quickly risk could be reversed). In order to take these effects into account in analyses, one could fit smoking status as a "time–updated covariate" in the Cox proportional hazards model. In the BRHS, this had the effect of increasing (marginally) the baseline hazard ratio for current cigarette smoking (from 1.78 to 1.92). Therefore, the estimates of the PARF for cigarette smoking presented in this chapter are likely to underestimate the true PARF,

though probably not by very much.

7.5.5 Conclusions: the validity of the "50% claim"

In stark contrast to the "only 50% claim" it is likely that at least 75–80% of first major CHD cases during middle–age may be attributed to the three strongest coronary risk factors: high serum total cholesterol, high blood pressure and cigarette smoking. However, this estimate will vary depending on how the low–risk group is identified (as the relationships between CHD risk and cholesterol and blood pressure are continuous and "threshold-free"). If all men in the BRHS had experienced the levels of risk of non–smokers, and had had the same cholesterol and blood pressure levels as those in the bottom fifths of the distribution, then an estimated 80% of all first major CHD events during middle–age would have been prevented. If the risks associated with all smoking (including previous smoking and passive smoking) could be eliminated, then up to 90% of all CHD events would have been prevented.

Table 7.1: Estimates of the regression dilution ratio (RDR) plus 95% confidence intervals taken over a four–year period for 259 men with no evidence of CHD at the 16–year screening

Risk factor	RDR	95% CI
Blood lipids		
Serum total cholesterol	0.72	(0.61,0.81)
(log) HDL cholesterol	0.86	(0.78,0.95)
(log) total:HDL cholesterol	0.76	(0.67,0.85)
Blood pressure		
Systolic blood pressure	0.62	(0.52,0.71)
Diastolic blood pressure	0.51	(0.42,0.61)
Mean arterial pressure	0.57	(0.48,0.67)

Table 7.2: Characteristics of 6,576 men with no baseline diagnosis or symptoms of CHD

	Observed value	Estimated usual value four years after baseline
Total cholesterol (mmol/L)		
Mean (SD)	6.3 (1.0)	6.3 (0.9)
10th centile	5.0	5.2
20th centile	5.4	5.5
HDL cholesterol (mmol/L)		
Median (IQR)	1.12 (0.97–1.29)	1.12 (0.98–1.29)
10th centile	0.85	0.86
20th centile	0.93	0.94
Ratio of total to HDL cholesterol		
Median (IQR)	5.5 (4.6–6.7)	5.5 (4.7–6.5)
10th centile	3.9	4.0
20th centile	4.4	4.5
Systolic blood pressure (mmHg)		
Mean (SD)	145 (21)	145 (16)
10th centile	121	124
20th centile	128	131
Diastolic blood pressure (mmHg)		
Mean (SD)	82 (13)	82 (10)
10th centile	66	70
20th centile	71	74
Mean arterial pressure (mmHg)		
Mean (SD)	103 (14)	103 (11)
10th centile	86	89
20th centile	91	94
Current cigarette smokers – no. (%)	2660 (41)	2660 (41)

SD = standard deviation; IQR = interquartile range

Table 7.3: The relative "informativeness" of models that adjust for different indices of blood cholesterol and blood pressure. Figures indicate reductions in deviance from the age–adjusted model and are expressed as percentages relative to the "informativeness" of the model that adjusts for total cholesterol and systolic blood pressure

	Blood pressure measure					
	Systolic		Diastolic		Mean arterial	
Blood lipid measure						
Total cholesterol	113.2	(100%)	99.2	(88%)	111.1	(98%)
(log) HDL cholesterol	73.8	(65%)	58.7	(52%)	71.7	(63%)
(log) total:HDL cholesterol	119.2	(105%)	103.0	(91%)	115.3	(102%)

Table 7.4: Predicted relative risk (RR) and population attributable risk fraction (PARF) for a range of threshold criteria before and after multivariate correction for regression dilution bias*

Definition of "low-risk" group	Threshold used	RR estimate (95% prediction interval) Before correction	RR estimate (95% prediction interval) After correction	PARF estimate (95% prediction interval) Before correction	PARF estimate (95% prediction interval) After correction
Usual serum total cholesterol					
< 5.0 mmol/L	Intervention	2.02 (1.68,2.40)	2.42 (1.70,3.82)	48% (38%,56%)	57% (39%,73%)
< 5.2 mmol/L	10th centile	1.97 (1.66,2.32)	2.31 (1.65,3.56)	45% (36%,53%)	54% (37%,70%)
< 5.5 mmol/L	20th centile	1.85 (1.58,2.14)	2.10 (1.56,3.06)	39% (30%,46%)	47% (31%,62%)
Usual systolic blood pressure					
< 124 mmHg	10th centile	1.78 (1.49,2.10)	2.58 (1.86,3.86)	40% (30%,48%)	59% (44%,72%)
< 131 mmHg	20th centile	1.70 (1.45,1.98)	2.30 (1.72,3.26)	35% (26%,43%)	51% (37%,64%)
< 140 mmHg	Intervention	1.64 (1.41,1.89)	2.10 (1.63,2.83)	28% (20%,35%)	41% (29%,53%)
Non-smoker		1.75 (1.47,2.15)	1.64 (1.36,2.03)	23% (16%,31%)	21% (14%,29%)
Serum total cholesterol < 5.5 mmol/L, systolic pressure < 131 mmHg and non-smoker	20th centile	3.45 (2.81,4.23)	5.24 (3.55,8.49)	70% (64%,75%)	80% (72%,87%)
Serum total cholesterol < 5.2 mmol/L, systolic pressure < 124 mmHg and non-smoker	10th centile	4.12 (3.27,5.19)	7.06 (4.47,12.55)	75% (69%,80%)	86% (79%,92%)
Serum total cholesterol < 5.0 mmol/L, systolic pressure < 140 mmHg and non-smoker	Intervention	3.60 (2.94,4.43)	5.33 (3.49,9.04)	72% (66%,77%)	81% (72%,88%)

* All RR and PARF estimates are adjusted for age, serum total cholesterol, systolic blood pressure and cigarette smoking status. The right hand columns also correct for regression dilution bias in both total cholesterol and systolic blood pressure

Table 7.5: Marginal relative risks and PARF estimates for each of the separate high–risk groups before and after correction for regression dilution bias

Risk groups defined by lowest usual decile			Before correction			After correction		
TC >5.2 mmol/L	SBP >124 mmHg	Smoker	$\hat{p}_j$	$\hat{RR}_j$	PARF$_j$	$\hat{p}_j$	$\hat{RR}_j$	PARF$_j$
×	×	×	0.01	1.00	-	0.01	1.00	-
√	×	×	0.08	1.96	2%	0.04	2.25	1%
×	√	×	0.07	1.78	1%	0.05	2.51	1%
×	×	√	0.01	1.79	0%	0.01	1.67	0%
√	√	×	0.44	3.64	28%	0.49	6.36	37%
√	×	√	0.05	3.43	3%	0.03	3.74	1%
×	√	√	0.05	3.13	2%	0.03	4.12	1%
√	√	√	0.29	6.30	38%	0.33	10.24	43%
Risk groups defined by lowest usual quintile			Before correction			After correction		
TC >5.5 mmol/L	SBP >131 mmHg	Smoker	$\hat{p}_j$	$\hat{RR}_j$	PARF$_j$	$\hat{p}_j$	$\hat{RR}_j$	PARF$_j$
×	×	×	0.04	1.00	-	0.03	1.00	-
√	×	×	0.11	1.86	3%	0.08	2.10	2%
×	√	×	0.10	1.71	2%	0.09	2.30	2%
×	×	√	0.03	1.77	1%	0.02	1.66	0%
√	√	×	0.35	3.26	23%	0.39	5.17	32%
√	×	√	0.07	3.24	5%	0.06	3.49	3%
×	√	√	0.07	2.98	4%	0.06	3.78	3%
√	√	√	0.23	5.63	32%	0.26	8.32	38%

TC = serum total cholesterol, SBP = diastolic blood pressure.

Table 7.6: Effect of adjustment for a range of other coronary risk factors on the relative risk and PARF (after correction for regression dilution bias) associated with low total cholesterol, low blood pressure and non-cigarette smoker

	Unadjusted*		Adjusted †	
Low-risk group	RR	PARF	RR	PARF
10th centiles	7.06	86%	5.94	83%
20th centiles	5.24	80%	4.51	77%
30th centiles	4.23	75%	3.71	72%

* as shown in Table 7.4.
†adjusted for usual physical activity level, usual alcohol intake, body mass index, history of diabetes, height, social class and town of residence.

Figure 7.1: Baseline distributions of total cholesterol, the logarithm of the ratio of total to HDL cholesterol, systolic blood pressure and diastolic blood pressure. Normal probability distributions have been superimposed.

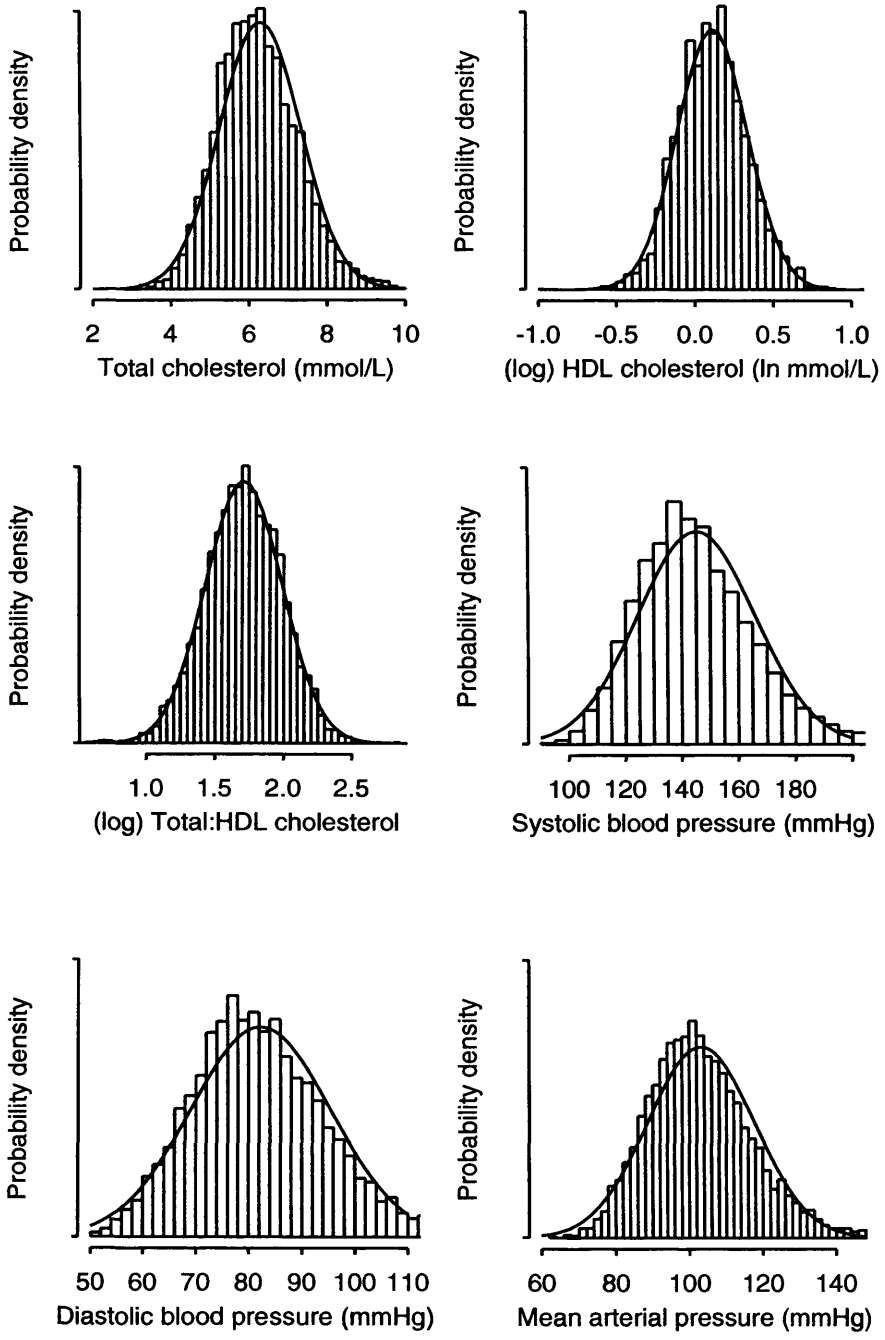
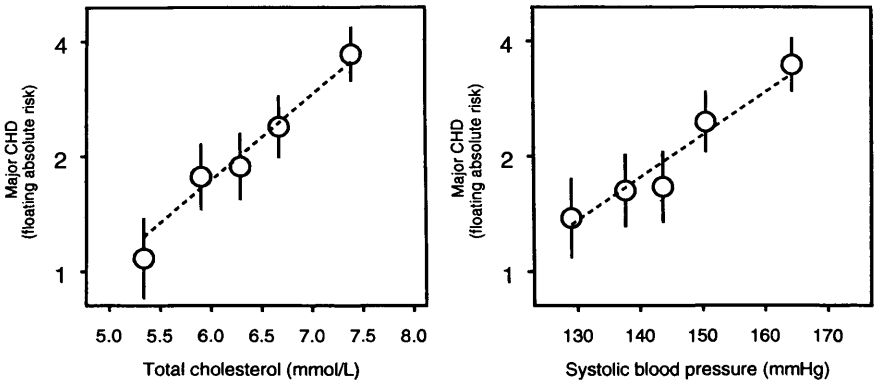


Figure 7.2: Floating absolute risks of major CHD within 10 years by usual serum total cholesterol (left) and usual systolic blood pressure (right). Estimates are presented for each of five equal sized groups and are adjusted for age, smoking status and, respectively, for systolic blood pressure and serum total cholesterol. Hazard ratio (HR) estimates adjusted for age, total cholesterol, systolic blood pressure and cigarette smoking status are shown below before and after multivariate correction for regression dilution bias



Risk factor	HR[1] (95% CI)	HR[2] (95% CI)
Age	1.07 (1.05,1.09)	1.05 (1.03,1.08)
Serum total cholesterol (1 mmol/l)	1.41 (1.30,1.53)	1.60 (1.35,1.96)
Systolic blood pressure (20 mmHg)	1.35 (1.24,1.47)	1.79 (1.47,2.18)
Cigarette smoking status		
Never/Ex	1.00 –	1.00 –
Current smoker	1.78 (1.47,2.16)	1.67 (1.15,2.36)

1. HR estimates from baseline data
2. HR estimates after multivariate correction for regression dilution bias in total cholesterol and systolic blood pressure.

Figure 7.3: Predicted relative risk and PARF before (thin line) and after (thick line) correction for regression dilution bias. The figures on the left shows these estimates for serum total cholesterol considered in isolation (high versus low); the figure on the right is for systolic blood pressure. In both graphs the lines correspond to (A) 20th centiles, (B) 10th centiles and (C) current clinical intervention thresholds

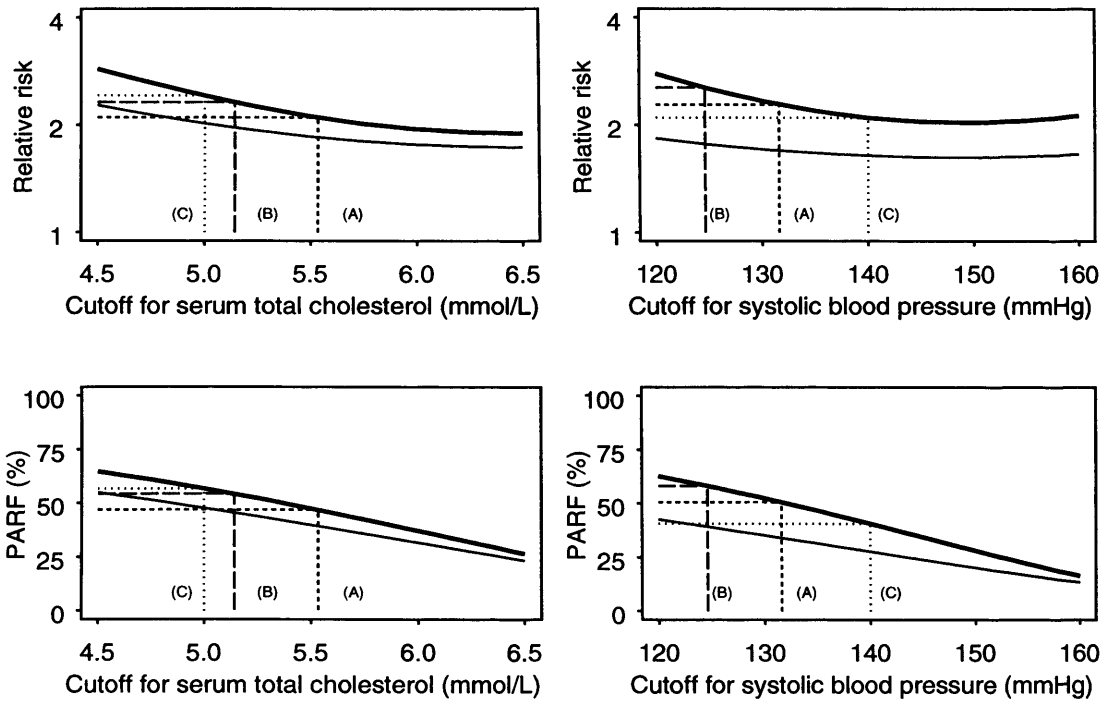


Figure 7.4: Predicted relative risk and PARF for serum total cholesterol, systolic blood pressure and cigarette smoking considered simultaneously before (thin line) and after (thick line) correction for regression dilution bias. The figures correspond to threshold criteria for systolic blood pressure based on (A) 20th centiles, (B) 10th centiles and (C) current clinical intervention thresholds. The dotted line in each panel corresponds to these criteria for serum total cholesterol.

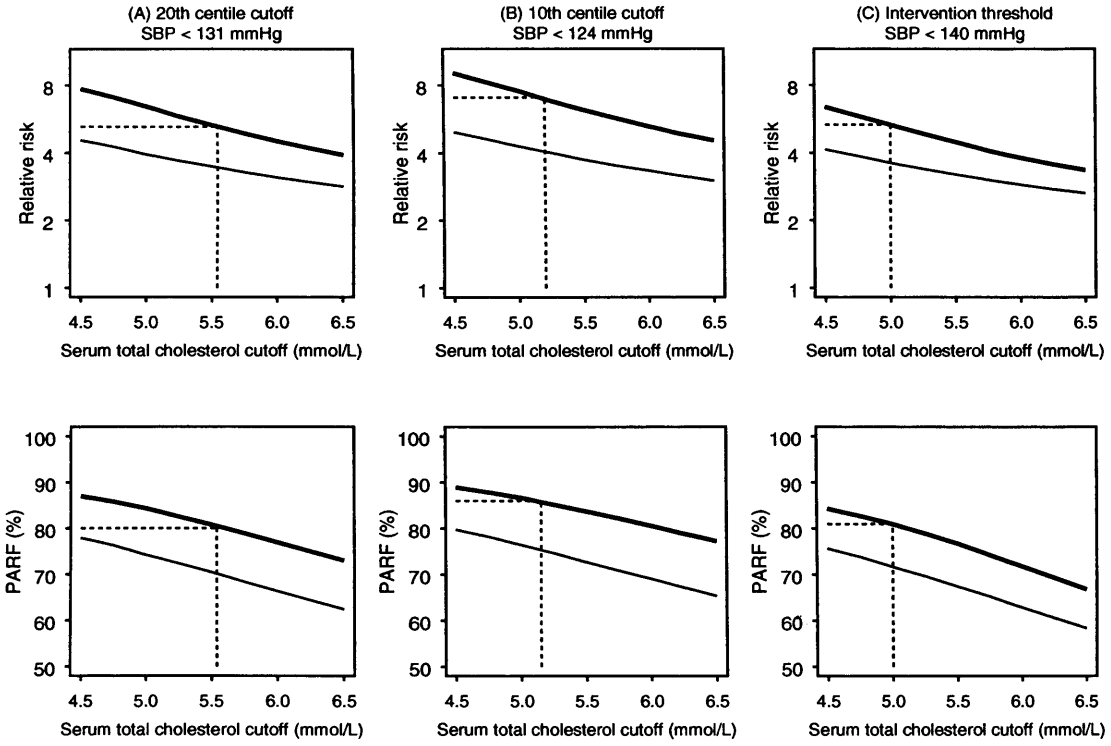
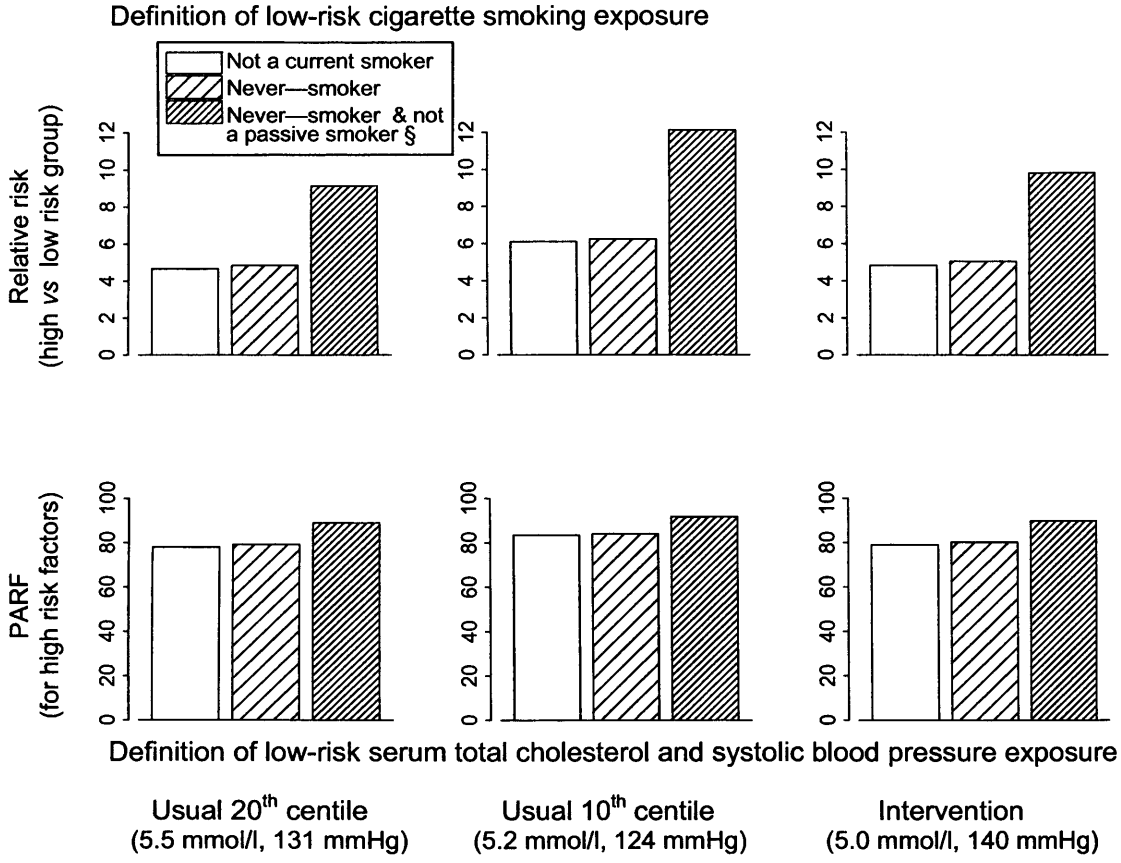


Figure 7.5: Effect of different threshold criteria for cigarette smoking on estimates of the relative risk and the PARF for high serum total cholesterol, high blood pressure and cigarette exposure. Analyses are based on 5,885 men with data on passive smoking.*



* corrected for regression dilution bias of serum total cholesterol and systolic blood pressure;
§ passive smoker defined as never regularly smoked cigarettes but with a cotinine level > 1.5 ng/ml.

Chapter 8

Strategies for the primary prevention of CHD

8.1 Summary

In this chapter, the long–term potential effectiveness of "population" approaches to the primary prevention of coronary heart disease, namely the population–wide reduction of important causal risk factors, are estimated before and after correction for regression dilution bias. These estimates are compared with the potential effectiveness of "high–risk" approaches to prevention, including the identification and treatment (with statins, aspirin, β–blockers and ACE inhibitors) of those men thought to be at greatest risk of developing CHD. The results indicate that even small downwards shifts in the population distributions of total cholesterol and blood pressure could have had a marked effect on reducing population levels of CHD. In men previously free from CHD, one in four of the major CHD events that occurred over the following ten years could have been prevented had average total cholesterol and systolic blood pressure levels been just 0.3 mmol/l (respectively 7 mmHg) lower (approximately 5% of the mean levels). Failure to take into account regression dilution bias would have led to this estimate being only one in five. In comparison, high–risk approaches were also potentially effective, but would have needed to have been used widely in order to have had a large impact on major CHD. Even if men with a predicted ten–year coronary event risk of 30% or more been identified and treated with a combination of a statin, aspirin, a β–blocker and an ACE inhibitor, only approximately one in nine of all major CHD events occurring within the following ten years would have

been prevented. The effects of regression dilution bias on the estimated effectiveness of the high–risk approach was negligible.

8.2 Introduction

8.2.1 Background

Two general strategies for the primary prevention of coronary heart disease are widely recognized – the "high–risk" approach, in which individuals at greatest risk of disease are identified and targeted for preventive treatment, and the "population" approach, in which population–wide changes in risk factors are made (see Figure 8.1).[48] The high–risk approach is the natural approach for medical practitioners who are concerned with the occurrence of disease in individuals. However, most CHD cases occur not amongst the small number of individuals at greatest risk, but amongst the much larger numbers of individuals at lower levels of absolute risk.[52] This is illustrated in Figure 8.2, where it can be seen that, amongst BRHS participants, individuals in the top fifth of the predicted risk distribution accounted for less than half of of all major CHD events that occurred over the following ten–years. It is the observation that risk is widely distributed in the population, and that CHD does not only affect those at greatest risk, that provides the basis for the great strength of the population strategy. Since the early description of these approaches for CHD prevention,[49;539] the potential impact of the strategies has changed. High risk approaches have been facilitated both by the availability of scoring systems to detect absolute CHD risk,[410] (rather than the traditional use of single risk factors) and by the advent of several treatments which produce marked and apparently independent reductions in CHD risk in high risk subjects.[51] However, it is also now recognised that the effectiveness of the population strategy has been underestimated by the failure to take account of regression dilution bias, so that relatively small reductions in the most important CHD risk factors (e.g. blood cholesterol and blood pressure) throughout the whole population could lead to unexpectedly large reductions in CHD.[49] Despite this, little attempt has so far been made to examine the potential impact of different high–risk strategies and population strategies, taking account of both advances in preventive treatments for CHD and the underestimation of population strategies introduced by regression dilution bias.

8.2.2 The Framingham equations and current prevention guidelines

The risk assessment methods most regularly used for identifying individuals at greatest risk of CHD are adapted from published equations derived from the Framingham Heart Study and the Framingham Offspring Study.[50] These studies followed 5,573 men and women aged 30–74 (who were initially free from cardiovascular disease) for 12 years, and derived risk prediction equations (based on age, gender, systolic blood pressure, cigarette smoking, the ratio of total to HDL cholesterol, diabetes and ECG evidence of left ventricular hypertrophy) for six outcomes: (1) myocardial infarction; (2) death from CHD; (3) all fatal and non-fatal CHD events; (4) stroke; (5) death from CVD; and (6) all fatal and non–fatal CVD events. Of these, the most commonly used equation is the "all CHD events" equation.

Current guidelines for primary prevention

In many European countries, the emphasis of current policy for the primary prevention of CHD is placed firmly with high–risk strategies, with little or no emphasis being placed on population–wide reduction of blood cholesterol and blood pressure.[4;408] In the United Kingdom, the National Service Framework (NSF) for Coronary Heart Disease[408] (published in March 2000) states that people without diagnosed CHD or other occlusive arterial disease but with a CHD risk greater than 30% over ten years (assessed using the Framingham CHD event equation[50]) should be offered:

- advice about how to stop smoking including advice on the use of nicotine replacement therapy.
- information about other modifiable risk factors and personalized advice about how they can be reduced (including advice about physical activity, diet, alcohol consumption, weight and diabetes).
- advice and treatment to maintain blood pressure below 140/85 mmHg.
- statins to lower serum cholesterol concentrations EITHER to less than 5 mmol/l (LDL-C to below 3 mmol/l) OR by 30% (whichever is greater).
- control of blood pressure and glucose in people who also have diabetes.

Similar high–risk strategies have been employed in Europe. The Second European Joint Task Force report on Coronary Prevention[4] published in 1998 stated that in primary prevention for individuals whose absolute risk of CHD is ≥20% over the next ten years (or will exceed 20% if projected to the age of 60), intensive risk factor modification is recommended including, where appropriate, a selective use of proven drug therapies (absolute risk was predicted using coronary risk charts based on the Framingham CHD equations). This report was revised in September 2003 by the Third European Joint Task Force Report on Cardiovascular Disease Prevention[417] which, for primary prevention, recommends targeting:

- individuals with multiple risk factors resulting in a ten-year risk of ≥5% now (or if extrapolated to age 60) for developing a **fatal CVD event,**
- individuals with markedly raised levels of single risk factors: total cholesterol ≥8 mmol/l, LDL cholesterol ≥6 mmol/l, blood pressure ≥180/110 mmHg,
- individuals with type 2 diabetes or type 1 diabetes with microalbuminuria,
- close relatives of patients with early onset atherosclerotic cardiovascular disease,
- close relatives of asymptomatic individuals at particularly high risk, or
- other individuals encountered in routine clinical practice.

These latest recommendations differ from previous guidelines in two ways. First, they are based on CVD risk, not CHD risk as previous guidelines have tended to use, and second, unlike previous guidelines they are not based on equations obtained from the Framingham study. Rather, a system of equations derived from the SCORE (Systematic COronary Risk Evaluation) project (derived using data from a number of different prospective studies including the BRHS) have been used to predict fatal CVD risk.[418] These equations have the advantage of being based on European populations with differing background risks of cardiovascular disease.

8.2.3 Preventive treatments and multiple risk factor reduction

The potential effectiveness of the high–risk approach to CHD prevention has been greatly influenced by the availability of several safe and effective risk reducing drugs, including

aspirin, β–blockers, diuretics, ACE inhibitors and statins, each of which has been shown to reduce CHD risk by 20% or more[84;118;359;540] (see Methods). Furthermore, the effects of these drugs have been shown to be fairly independent of one another, so that by taking multiple drugs, large reductions in CHD risk may be achievable.[51] Recently, several authors have commented on this possibility, and in particular the possibility of combining these separate components into a single pill.[51;425] Indeed, it has been claimed that a combination pill consisting of aspirin, a statin, three blood pressure lowering drugs at half dose, and folic acid (the "polypill") may be able to reduce CHD risk in individuals by as much as 88%.[425]

8.2.4 Objectives

In this chapter, the potential effectiveness of population strategies (directed at the control of blood pressure and blood cholesterol levels in the population), and the potential effectiveness of different high–risk strategies to the primary prevention of CHD (including both policies directed to the measurement and control of individual risk factors, particularly cholesterol and blood pressure, and those based on the identification and management of high overall CHD risk) are examined. In particular, relationships between the established risk factors and major CHD risk in the first ten years of follow–up in the BRHS (before and after correction for regression dilution bias) are used to predict the proportion of all major CHD events that may have been prevented had various prevention policies been implemented.

8.3 Methods

Using a combination of data from the British Regional Heart Study and estimates of relative risk reductions from meta–analyses of randomised controlled trials (where possible), the likely influences of population and high–risk strategies on first occurrences of major CHD events in middle–aged British men are assessed. As in the previous chapter, analyses are restricted to individuals with no baseline diagnosis or symptoms of CHD and only first major CHD events over the first ten years of follow–up are used in analyses. In this chapter, because of the specific interest in the potential role of preventive strategies, analyses are further restricted to men not receiving lipid lowering or blood pressure lowering drugs

at baseline.

8.3.1 Population approaches to prevention

Three specific population approaches to prevention were considered:

1. Population–wide reduction in mean cholesterol.
2. Population–wide reduction in mean blood pressure.
3. Population–wide reductions in both mean cholesterol and mean blood pressure.

The effectiveness of these approaches were estimated as the epidemiologically expected reductions in major CHD risk corresponding to having certain lifetime lower measurement levels. Thus, it was implicitly assumed that the observed statistical associations between these exposures and major CHD risk were of a causal nature, reflecting the true differences in risk caused by long term differences in blood cholesterol and blood pressure level. Care was therefore taken to adjust for any potential confounding risk factors; analyses were adjusted for a range of other major risk factors as well as socio–demographic factors (see statistical methods; § 8.3.3). For each of the population approaches considered, the estimated effectiveness relates to the effects of producing a "downwards shift" in the population distributions of cholesterol and blood pressure (see Figure 8.1), not a proportional reduction in each individual's measurement level. However, for presentation purposes, these downwards shifts are expressed as percentages of the mean risk factor level observed in the population.

8.3.2 High–risk approaches to prevention

A variety of different high–risk prevention strategies were considered including treatment with single drugs and treatment with multiple combined drugs. These included:

1. Identification and management of individual risk factors:
   (a) set treatment threshold level for blood cholesterol and treat with a statin;
   (b) set treatment threshold level for blood pressure and treat with a β–blocker or diuretic.

2. Identify treatment threshold level of 10–year Framingham risk (as recommended in the UK at a $\geq$ 30% level)[408] and treat with:

   (a) a statin;

   (b) a β–blocker or diuretic;

   (c) a combination of aspirin, a β–blocker or diuretic, an ACE inhibitor and a statin.

Estimation of drug effects

Estimates of relative risk reductions for each class of drug were taken from relevant randomised controlled trials or (where possible) meta–analyses of such trials. It was assumed that blood cholesterol reduction with statins reduced the risk of MI by 31%,[84] and that blood pressure reduction with first line antihypertensive drugs (a diuretic or a β–blocker) reduced MI risk by 18%.[540] Among "high–risk" individuals, it was also assumed that aspirin reduced the risk of MI by 26%[359] and ACE–inhibitors reduced MI risk by 20%.[118] Treatment effects were assumed to be multiplicative so that the combined relative risk reduction from taking aspirin, statins, ACE–inhibitors and β–blockers/diuretics was 67% (100% $\times$ [1 - 0.74 (aspirin) $\times$ 0.69 (statins) $\times$ 0.80 (ACE inhibitors) $\times$ 0.82 (β–blockers/diuretics)]).[541] In a subsidiary analysis, the potential effects of a prevention approach based on age and the use of a combination of treatments including aspirin, a β–blocker or diuretic, an ACE inhibitor and a statin was examined. This analysis was performed in order to provide a means of discussing the recent paper from Wald and Law, which suggested that CHD could potentially be reduced by up to 88% by treating all individuals over the age of 55 with multiple drugs (the "polypill").[425] [i]

Identification of high–risk individuals

The Framingham "CHD event" equation was used to estimate an individual's overall absolute risk of CHD risk over the first ten years of follow–up from their baseline levels. This method was used in preference to other approaches because, at the time of writing, the Framingham equations provide the most common method of estimating CHD risk,

[i] NB/ The difference between the 67% relative risk reduction assumed in this chapter and the 88% reduction estimated by Wald and Law in their "polypill' paper is due mainly to their assertion that statin therapy alone could reduce CHD risk by as much as 61% (their estimate was based on cohort studies rather than clinical trials; see section 8.5.3 for discussion), as well as the additional component of folic acid in the polypill (assumed to reduce risk by 16%) which does not feature in the combination treatment described in this chapter.

particularly in the United Kingdom. Though there is growing evidence that Framingham equations tend to overestimate true risk in European populations[419–422;424;542] (including evidence from the BRHS),[423] the original equations are used throughout this chapter to reflect current guidelines. The implications of using a score that may well overestimate the true risk of CHD in British men are discussed in section 8.5.

8.3.3 Statistical methods

The regression model and the estimates of the RDR

The association between baseline risk exposures and ten–year major CHD risk was assessed using Cox proportional hazards regression; analyses were adjusted for age, blood cholesterol, blood pressure, cigarette smoking status, body mass index, physical activity, alcohol intake, history of diabetes and area of residence (the South, Midlands & Wales, the North, Scotland). As in the previous chapter, the relative abilities of different indices of blood cholesterol (total cholesterol, HDL cholesterol and the ratio of total to HDL cholesterol) and blood pressure (systolic, diastolic and mean arterial) to predict major CHD risk was assessed through examination of its χ^2 likelihood ratio statistic in the fully adjusted model. These analyses are repeated in this chapter because a wider range of risk exposures are considered in the model. Furthermore, the sample differs slightly from that used in chapter 6 as it excludes men taking drugs to lower cholesterol or blood pressure at baseline.

Estimates of the regression dilution ratio were taken over the four year interval between the 16– and 20–year screenings in Dewsbury and Maidstone. Analyses were restricted to men with no previous evidence of CHD (at the 16–year screening) and not taking lipid lowering or blood pressure lowering drugs at either the 16– or 20– year screenings. These estimates were used to examine true associations over the first ten years of follow–up from observed "baseline" associations (regression calibration (see equation 3.20) was used to estimate expected usual exposure levels and true regression coefficients).[26]

Predicting the effectiveness of the prevention strategies

Having determined the "most informative" blood cholesterol and blood pressure indices for CHD risk prediction purposes (and having corrected the regression coefficients for regression dilution bias), the potential effectiveness of each of the "high–risk" prevention

strategies was assessed using the regression results to predict absolute CHD risk corresponding to a certain set of covariates (using the recalibrated blood pressure and blood cholesterol measurements). When predictions are made on the sample from which the prediction tool is derived, estimates of risk differences may be biased, often seriously.[543] Therefore, predicted risks were obtained using the "jack–knife" technique (described in section 3.4.4) which eliminates such bias from being introduced.[485] The mean of these predicted risks provides an estimate of the expected ten–year absolute CHD risk in the population prior to implementation of the prevention strategy (which should be close to the observed CHD risk). Individuals whose observed risk exposure levels were sufficiently high to warrant preventive treatment (i.e. the high risk group) subsequently had their predicted risks recalculated to take into account the effects of treatment. The mean predicted risk after implementation of the strategy was then calculated, allowing calculation of the expected reduction in major CHD risk due to the high–risk strategy. For the "population strategies", the expected reduction in major CHD over 10 years was estimated by comparing the predicted CHD risk in the observed sample with the predicted risks for the same sample following absolute reductions in each individual's blood cholesterol and blood pressure level. For these approaches, the reductions in major CHD correspond to reductions that would be expected if the sample had had lifetime lower blood pressure and blood cholesterol levels.

8.4 Results

Of the 6,576 men with no baseline evidence of CHD, 220 were receiving blood pressure or lipid–lowering lowering drugs and were excluded from these analyses. Of the remaining men, 6,011 (94.6%) had complete risk factor data. The baseline characteristics of these men are displayed in Table 8.1. Over the following ten years of follow–up, 371 men (6.2%) had a major CHD event. The "relative informativeness" of different blood cholesterol and blood pressure indices at predicting CHD risk was assessed by examining the likelihood ratio χ^2 statistics in the fully adjusted model. As was previously observed, systolic pressure was more informative than diastolic pressure and total cholesterol was more informative than HDL cholesterol. Mean arterial pressure and the ratio of total to HDL cholesterol were at least as predictive as systolic pressure and total cholesterol, but the latter indices

were again preferred for simplicity.

Repeat blood pressure and blood cholesterol measurements over four years (between 16 and 20 years) were available for 165 men with no previous evidence of CHD and not on treatment to lower blood pressure or blood cholesterol at either the 16– or 20–year examinations. Regression dilution ratio estimates of 0.79 (95% CI 0.69 to 0.89) for total cholesterol and 0.75 for systolic blood pressure (95% CI 0.63 to 0.88) were observed.

8.4.1 Effectiveness of population strategies for prevention

Using the estimated relative hazards for total cholesterol and systolic pressure (after adjustment for a range of other factors and for regression dilution bias) the predicted effectiveness of each of the "population" approaches to prevention was estimated. This is shown in Figure 8.4 and Table 8.2. If the total cholesterol levels of all men had been 0.3 mmol/l lower and the systolic blood pressure levels been 7 mmHg lower (corresponding to reductions in average blood cholesterol and blood pressure of 5%), then it is estimated that there would have been 24% fewer first major CHD events during the first ten years. Had these reductions been twice as great (so that total cholesterol had been 0.6 mmol/l lower and systolic blood pressure been 14 mmHg lower in all men), then there would have been an expected 42% fewer major CHD events.

The effects of regression dilution bias on the estimated effectiveness of the population approach are shown in Table 8.3. It can be seen that for the population approaches considered, the true effect sizes are between 20 and 30 percent greater than the uncorrected estimates (depending on the size of the population shifts), so that if the "naïve" estimate were a 40 percent reduction in CHD, the true estimate would be nearer 50 percent.

8.4.2 Effectiveness of high–risk strategies for prevention

Table 8.2 shows the estimated effectiveness of each of the high–risk policies at specific "treatment thresholds" and Figure 8.3 the relations between the thresholds, effectiveness and the proportion of the population treated under the strategy. As the threshold for treatment reduces (i.e. as the proportion of the population treated increases) the estimated reduction in major CHD events in the population increases. For a given intervention, the effectiveness of identification based on overall risk (through the calculation of a Framingham risk score) is generally greater than that based on identification of single risk

factors, and becomes more so as thresholds fall. Multiple interventions have considerably greater benefits in terms of prevention than intervention based only on blood cholesterol or blood pressure. However, even with multiple drug treatment, the predicted reduction in first occurrences of major CHD following preventive treatment at a threshold based on a Framingham risk of $\geq$30% (as is currently recommended in the UK), was only 11%. This increased to 33% when the Framingham threshold was reduced to a 10-year risk of $\geq$20% (as recommended by the European Joint Task Force on Coronary Prevention), and 48% when the threshold was reduced to $\geq$15%. At these lower thresholds, respectively one quarter and one half of the population without symptomatic CHD would be receiving multiple preventive treatment. Regression dilution bias had virtually no effect on the estimated effectiveness of the high–risk approach, as treatment effects were taken from randomised controlled trials.

Treatment based on age criterion alone

For the 371 men who experienced a first major CHD event over the ten–year follow–up period, 241 of them (65.0%), were aged 55 or over at the time of the event. If a prevention policy were introduced whereby men received the four–drug intervention when they reached the age of 55, then 161 of these first events (241 x 0.67) could potentially have been prevented. Therefore, approximately 43% of all first major CHD events over the following ten years (161/371) may have been prevented through implementation of this particular high–risk policy (assuming 100% prescription rates and adherence levels as high as those observed in the clinical trials). Reducing the age threshold to 50 would have increased this proportion to 59% ((325 x 0.67) / 371).

8.5 Discussion

8.5.1 Interpretation of findings

Taking regression dilution bias of blood cholesterol and blood pressure into account, the potential effectiveness of a variety of high–risk and population strategies for the primary prevention of CHD was estimated. The results indicate that in order to have a substantial effect on CHD rates, high–risk multiple–intervention primary prevention policies would need to be used widely – at a level below the 3% predicted risk per annum recommended

in the UK,[408] and possibly even below the 2% predicted risk per annum recommended by the European Joint Task Force on Coronary Prevention.[4] In comparison, relatively small absolute reductions in the population levels of two key risk factors (blood cholesterol and blood pressure) would potentially lead to large reductions in population levels of major CHD.

In these analyses, correction has been made for the effects of regression dilution bias (the underestimation of relationships between usual risk factor levels and disease risk caused by within–person variability). Though the high–risk approaches were robust to these effects (as treatment effects were taken from trials), the effects of regression dilution bias on the estimated effectiveness of the population approaches were marked (Table 8.3). This is because the true size of the shift in the exposure distribution relative to the exposure variance is greater than would be estimated if within–person variability were not taken into account. Put differently, a naïve analysis of the effect on CHD of reducing cholesterol in the population by a certain level would produce an answer that, in reality, corresponds to a smaller change in population cholesterol levels. In such analyses, it is therefore crucially important to take regression dilution bias into account, as failure to do so would almost certainly lead to underestimation of the effectiveness of the strategy.

8.5.2 Effectiveness and feasibility of the population approach

The effectiveness of the population approach to prevention depends critically on the size of population–wide changes that could realistically be achieved in practice. The population–wide reductions in total cholesterol and blood pressure assumed in this paper (Table 8.2) are relatively modest and are consistent with reductions that may be achievable though changes in diet.[104;127;426–430;544] In North Karelia, Finland, a community based cardiovascular disease prevention project (initiated in 1972 and followed up to 1997) has resulted in population cholesterol levels being reduced by 18% in both men and women and blood pressure being reduced by 8% in men and 13% in women.[426] These changes in the population distributions of blood cholesterol and blood pressure were estimated to explain most of the 55% reduction in CHD mortality observed during that period.[544] Had mean cholesterol and blood pressure levels been 18% (respectively 8%) lower in the BRHS, it is predicted that major CHD would have been 51% lower. The consistency between this figure and the figure actually observed in Finland following these same changes in

cholesterol and blood pressure is remarkable. Large population–wide changes in cholesterol have also been observed elsewhere. A study of the effects of a population–wide intervention programme in Mauritius found that between 1987 and 1992, average total cholesterol fell from 5.5 mmol/l to 4.7 mmol/l following a change in the island's supply of cooking oil from palm to soy bean oil and the implementation of an intervention programme aimed at the promotion of a healthy lifestyle.[427] In Western countries, meta–analysis of metabolic ward studies has indicated that, irrespective of age, sex or body weight, replacing saturated fats with complex carbohydrates for around 10% of dietary calories reduces serum total cholesterol by around 0.5 mmol/l (four–fifths of this reduction being in LDL cholesterol), and that further reductions are possible by replacing complex carbohydrates with polyunsaturated fats and avoiding dietary cholesterol intake.[104] For typical British diets, the results indicate that if 60% of saturated fats could be replaced by other fats and 60% of dietary cholesterol could be avoided, mean serum total cholesterol in the population would fall by approximately 0.8 mmol/l (that is, by 10–15%). Results from the Health Survey for England show that in middle–aged men and women, mean total cholesterol has fallen by between 0.4 and 0.6 mmol/l between 1994 and 1998,[545] although this figure could be subject to bias because of a change in the laboratory methods used over this period.

Similarly, population–wide reduction in the amount of dietary salt could significantly reduce population blood pressure levels. For instance, in the Mauritius study mean systolic pressure fell by approximately 5 mmHg and mean diastolic pressure by approximately 2.5 mmHg over the five year study period,[427] while in a community study of salt restriction in Portugal, systolic and diastolic pressure both fell by an average of 5 mmHg within two years.[430] A meta–analysis published in 1991 of 45 salt reduction trials estimated that reducing daily sodium intake by 50 mmol (about 3 grams of salt) would, after only a few weeks, reduce systolic blood pressure by an average of 5 mmHg and diastolic by around half as much.[127] Interestingly, the authors estimated that reducing salt intake by this amount throughout the entire population would reduce the incidence of CHD by 16%, slightly more than our estimate (12% reduction in major CHD following a 7 mmHg (5%) reduction in SBP; see Table 8.2). The authors proceeded to claim that reduction also in the amount of salt added to processed foods could lower population blood pressure by around twice as much (10 mmHg systolic and 5 mmHg diastolic). However, clinical advice to restrict salt intake is likely to be less effective, because of low levels of compliance

as well as the fairly limited control an individual actually has on determining their salt intake. In a systematic review of trials examining the effects of advice to reduce salt intake carried out in 2002, only small reductions in blood pressure were observed.[546] The authors suggested that intensive behavioural interventions were unsuited to population primary prevention programmes. Changes in population levels of blood pressure are achievable however; secular changes have been observed over relatively short periods of time. In a study of students entering Glasgow university between 1948 and 1968, mean systolic blood pressure for male students born after 1945 was 10 mmHg lower than for male students born before 1929; diastolic pressure was 5 mmHg lower on average.[547] More recently, trends of a similar size, independent of antihypertensive treatment, have been reported in the Health Survey for England, where systolic blood pressure in men aged 55 to 74 has fallen by 2–3 mmHg over a four year period (1994 to 1998).[545]

Effects of other risk factors

In this analysis, attention has been focussed on cholesterol, blood pressure and pharmacological interventions and no account has been taken of the important additional contribution made to CHD risk by cigarette smoking. Taking account of cigarette smoking would made a notable additional contribution to the effectiveness of both population and high–risk approaches (for instance, approximately one third of the falls in CHD mortality in Scotland over the last two decades has been estimated to be due to reductions in cigarette smoking).[548] However, the balance of potential effectiveness of the two strategies is unlikely to be affected by taking cigarette smoking into account. The balance of effectiveness between the two strategies may be affected in other ways however. The high–risk approaches considered in this chapter were primarily focused on reducing CHD risk through reducing blood cholesterol and blood pressure (and, in the case of aspirin, clotting processes). For the purpose of comparison, population approaches were selected that aimed to cause a reduction in blood pressure and blood cholesterol at the population level. In reality however, it is likely that strategies directed towards lowering blood pressure and blood cholesterol in the population will have additional favourable benefits on other coronary risk factors, such as body mass index and, potentially, physical activity, the effects of which have not been taken into account in these analyses. In addition, population approaches are likely to offer benefits for secondary prevention, both by reducing the risk of subsequent CHD

events in individuals already with clinically manifest CHD (which would be unaffected by high–risk primary prevention approaches) and by reducing the overall burden of CHD in the population (which in turn could enable secondary prevention strategies to be better focused).

8.5.3 Effectiveness of the high–risk approach

Assumed treatment effects

The validity of the estimates for the high–risk approaches depends on the treatment effects assumed and the appropriateness of the strategies. The effects of statins, aspirin, and first line blood pressure lowering drugs were taken from meta–analyses of randomised controlled trials,[84;359;540] while the effects of ACE inhibitors were estimated from a large randomised controlled trial of these agents[118]. These estimates were used in preference to estimates from cohort studies[425] because while cohort studies allow estimation of the effects of differences in risk due to long–term differences in risk exposure levels, clinical trials show the extent to which these epidemiological associations are reversible through treatment. The clinical trial estimates also take non–adherence into account as they are obtained from "intention–to–treat" analyses, though they may still overestimate the true efficacy of the drugs in routine practice because: (1) trials often exclude those with poor adherence in "run–in" phases; (2) supervision is often more systematic than in usual care; and (3) adherence is likely to decrease over the longer term.[549;550] In addition, since the treatment effects were generally obtained from studies of "high–risk" men (including men with previous CHD), by applying the results to men without previous CHD, the effectiveness of the high–risk approach may have been overestimated. This may be particularly true for ACE inhibitors, for which evidence of effectiveness is substantially based on subjects with established CHD.[118] For statins and aspirin, this assumption is more clearly valid, as relative risk reductions are reasonably stable across a wide range of risk groups.[78;359] Furthermore, by assuming that the treatment effects were multiplicative, the combined effects of taking all four drugs may have been overestimated (for example ACE inhibitors may be less effective when used in combination with aspirin).[551] However, by combining different combinations of drugs (including multiple low dose drugs), greater reductions in CHD risk may be possible than assumed in this chapter,[122] though even if this were the case, it is unlikely that the estimates would be greatly affected (if the true relative

risk reduction of the combined pill was 85%, for instance, treatment of individuals with a $\geq$ 30% Framingham risk would reduce major CHD by 14%, compared with 11% if the four drugs described in this chapter were used).

Use of the Framingham risk equations to identify high–risk individuals

In order to reflect current practice in the United Kingdom, the analyses in this chapter have identified "high–risk" individuals based on the Framingham CHD event equation (published by Anderson and colleagues in 1991).[50] This equation allows individuals to be ranked in order of risk so that a particular group (say those with a ten–year risk of $\geq$30%) may be identified and offered treatment. Recently however, many authors have found that the Framingham equations tend to overestimate absolute risk in European populations.[419–422;424;542] Indeed, recent evidence from the BRHS has shown that the Framingham equations overestimate true CHD risk in British men by around 50%.[423] What effect, therefore, would a recalibrated Framingham risk score have on the estimated effectiveness of the high–risk approach? If the predicted CHD risks from the Framingham equation were all divided by 1.5 (in order to take into account this overestimation) then the number of individuals meeting any specific "threshold" level for treatment (e.g. $\geq$30% or $\geq$20%) would be greatly reduced. Therefore the number of individuals treated would decrease and the estimated effectiveness of the high–risk approach would be reduced. In the British Regional Heart Study, of the 6,011 men used in the analyses presented in this chapter, 2830 (47%) had a predicted Framingham 10–year CHD event risk of 15% or more. However, if each of these scores were divided by 1.5 in an attempt to recalibrate the equation, then only 1143 men (19% of the whole population) would still have a predicted risk of $\geq$15% (i.e. men whose predicted risk was originally $\geq$22.5%). This would result in the effectiveness of the "high–risk approach" being substantially reduced (from a 48% reduction in major CHD to a 26% reduction). In practice however, it is likely that in Western societies (where most individuals are at least at moderate levels of risk) the thresholds used to define the high–risk group may be determined by the availability of resources, rather than by the level deemed worthy of intervention. If so, it is the ability of a risk equation to "rank" individuals in order of risk that is likely to be most relevant.

8.5.4 Conclusions: approaches to CHD prevention

High–risk approaches to prevention are effective (reducing the risk of CHD in individuals by 70% or more for multiple combined drugs), but unless they are used widely, their impact on the level of first cases of CHD in the population is likely to be fairly small. However, small downwards shifts in the risk factor distributions of the key coronary risk factors (such as blood cholesterol and blood pressure) could have a large impact on population levels of disease in the long–term (perhaps even greater than estimated in this chapter), despite the small absolute risk reductions experienced by each individual. The estimated effectiveness of the population approach is particularly prone to underestimation due to regression dilution bias, and it is perhaps partially for this reason that its potential for reducing CHD has gone largely unrecognised. Population approaches targeted through changes in the national diet may be possible to implement with little noticeable effect to the individual and should eventually lead to a decreased pool of middle–aged people requiring drug treatment.

Table 8.1: Baseline characteristics of 6,011 men with no baseline evidence of CHD and not receiving blood pressure lowering or lipid–lowering drugs at baseline. Data correspond to mean (SD) unless otherwise stated

Baseline risk factor	Observed value
Age (years)	49.8 (5.8)
Serum total cholesterol (mmol/L)	6.3 (1.0)
Total:HDL cholesterol*	5.5 (4.6 - 6.7)
Systolic blood pressure (mmHg)	145 (20)
Diastolic blood pressure (mmHg)	82 (13)
Body mass index (kg/m^2)	25.4 (3.2)
Current cigarette smokers – no. (%)	2438 (41)
At least moderately active – no. (%)	2384 (40)
Heavy drinkers – no. (%)	630 (11)
History of diabetes – no. (%)	68 (1)

* Geometric mean presented (interquartile range)

Table 8.2: A comparison of approaches to the primary prevention of CHD

Prevention approach		RRR	Predicted reduction in major CVD		
"High–risk" approach	Management		Group identified for treatment		
			Top 10%	Top 20%	Top 30%
Treat high total cholesterol	Statin	31%	6%	10%	13%
Treat high blood pressure	β–blocker/diuretic	18%	4%	6%	8%
Treat high total cholesterol	Four drug combination	67%	14%	22%	29%
Treat high blood pressure			16%	24%	30%
Treat high overall absolute risk			16%	27%	36%
			Framingham 10-year CHD risk		
			$\geq 30\%$	$\geq 20\%$	$\geq 15\%$
Treat high overall absolute risk	Statin	31%	5%	15%	22%
	β-blocker/diuretic	18%	3%	9%	13%
	Four drug combination	67%	11%	33%	48%
"Population" approach			Shift the risk factor distribution by		
			5%	10%	15%
Reduce mean total cholesterol in the population			13%	24%	34%
Reduce mean blood pressure in the population			12%	23%	33%
Reduce mean total cholesterol and blood pressure in the population			24%	42%	56%

RRR = relative risk reduction (for the high–risk approach); four drug combination = aspirin, statins, β–blockers/diuretcis and ACE inhibitors

Table 8.3: Effect of regression dilution bias on the estimated effectiveness of the population approach. The figures shown are the amounts by which the uncorrected estimates of effectiveness should be multiplied by to obtain the true estimates.

	Population reduction in systolic pressure			
	0 mmHg	7 mmHg	15 mmHg	22 mmHg
Population reduction in total cholesterol				
0 mmol/L	NA	1.31	1.29	1.27
0.3 mmol/L	1.27	1.26	1.26	1.24
0.6 mmol/L	1.24	1.24	1.23	1.21
0.9 mmol/L	1.23	1.22	1.21	1.20

Figure 8.1: The "population" approach to CHD prevention aims to cause downwards shifts in the most important causal risk factors throughout the whole population.

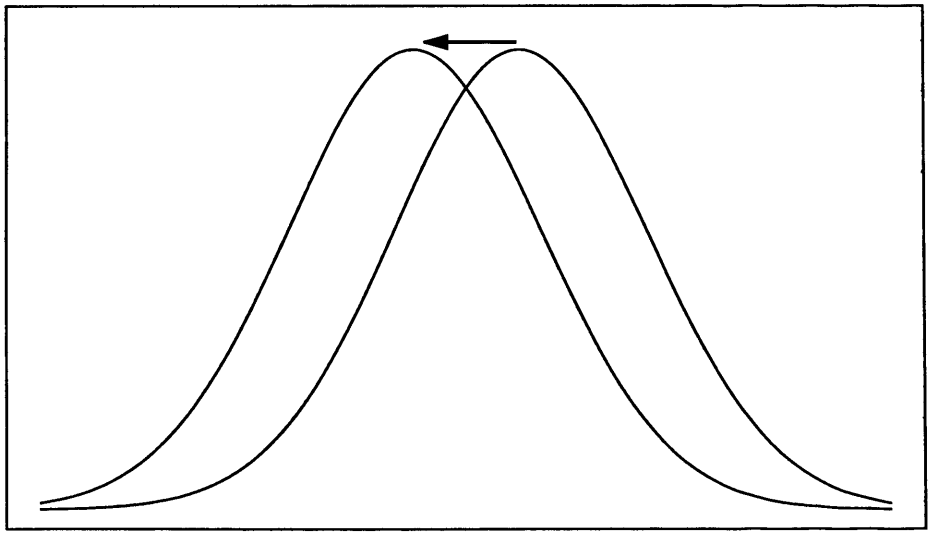
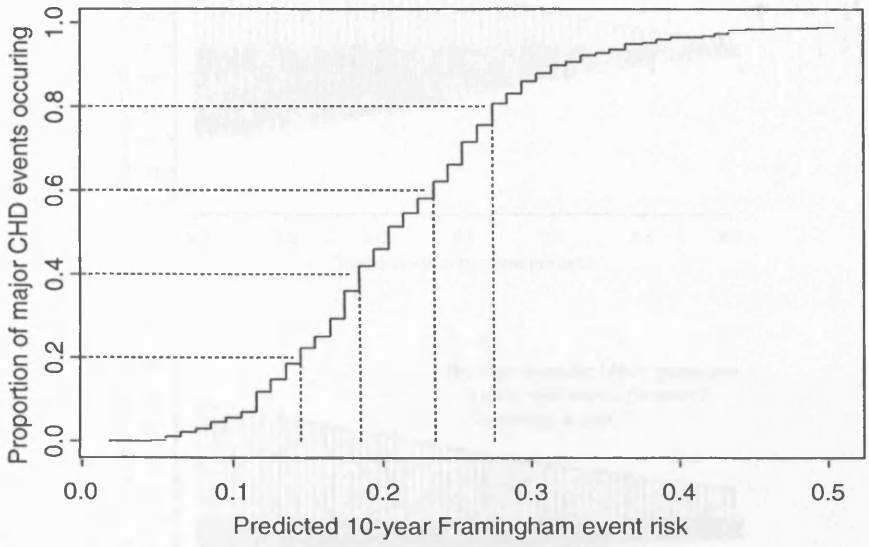


Figure 8.2: Proportion of all major CHD events occurring in the BRHS (within first 10 years) by Framingham predicted CHD event risk.



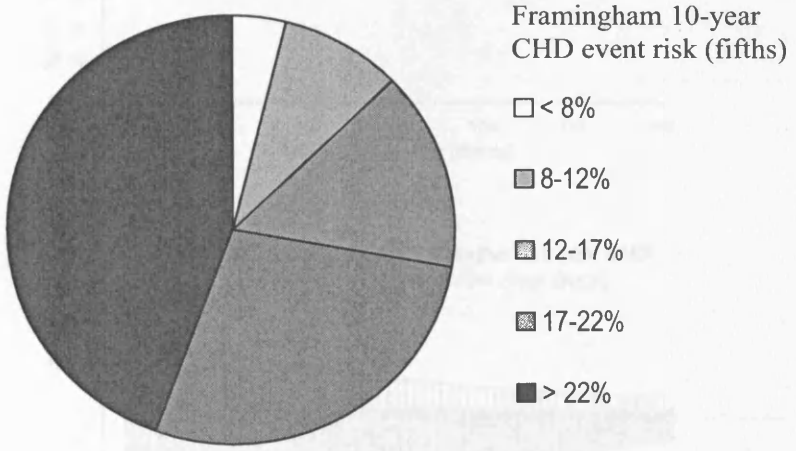


Figure 8.3: Predicted effectiveness of three different "high–risk" approaches to CHD prevention as a function of the "threshold" used to define the high–risk group. In each case the bars above the axis indicate the proportion of the population treated and the bars below the axis show the expected reductions in major CHD as a result of the policy

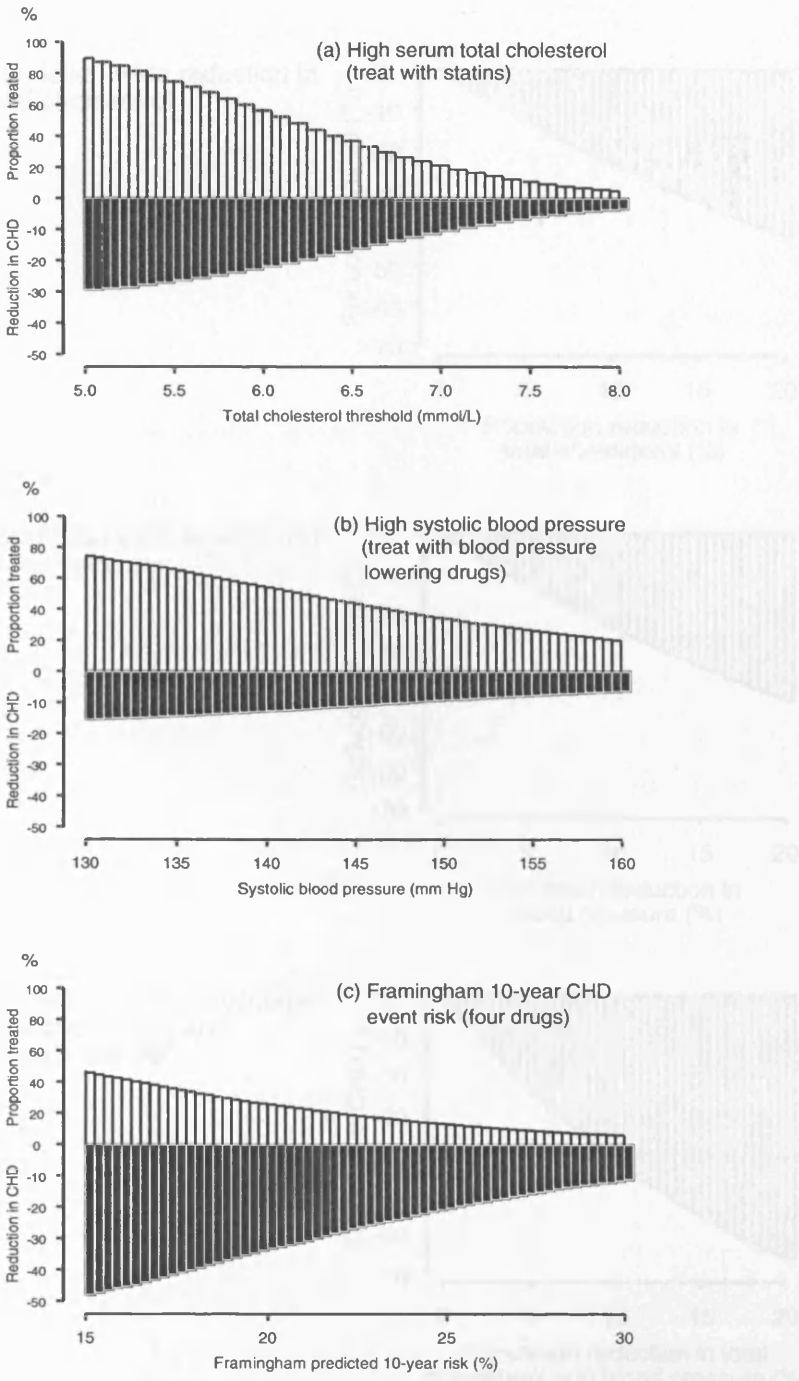
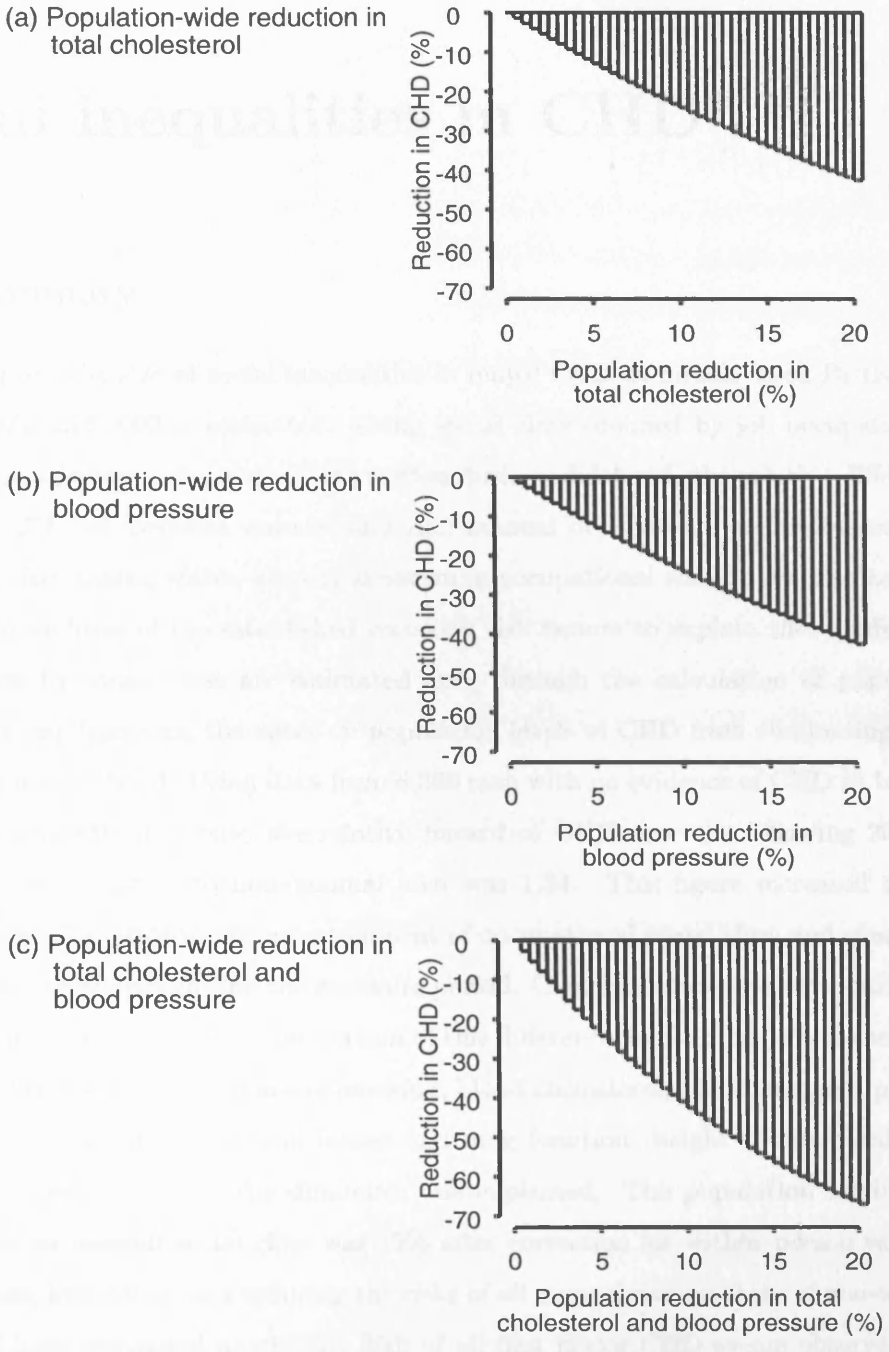


Figure 8.4: Predicted effectiveness of three "population" approaches to CHD prevention. In each case the bars indicate the expected reduction in major CHD following long–term reductions in: (a) blood cholesterol; (b) blood pressure; and (c) both blood cholesterol and blood pressure



Chapter 9

Social inequalities in CHD

9.1 Summary

In this chapter, the size of social inequalities in major CHD in middle–aged British men between 1980 and 2000 is estimated. Using social class (defined by job occupation) as a marker for a person's socioeconomic position during adulthood, the relative differences in major CHD risk between manual and non–manual occupations are evaluated both before and after taking within–person variation in occupational social class into account. The relative abilities of the established coronary risk factors to explain these differences in CHD risk by social class are estimated and, through the calculation of population attributable risk fractions, the effect on population levels of CHD from eliminating social inequalities is calculated. Using data from 6,386 men with no evidence of CHD at baseline and not in the Armed Forces, the relative hazard of CHD over the following 20 years for manual men relative to non–manual men was 1.34. This figure increased to 1.41 after correction for error in the ascertainment of occupational social class and changes in occupational social class during the exposure period. Cigarette smoking status during the study accounted for the greatest proportion of this difference, explaining 38% of the excess risks of manual men. When cigarette smoking, blood cholesterol, blood pressure, physical activity, body mass index, alcohol intake and lung function (height standardised) were considered together, 41% of the difference was explained. The population attributable risk fraction for manual social class was 19% after correction for within–person variation in social class, indicating that reducing the risks of all manual men to those of non–manual men would have prevented nearly one fifth of all first major CHD events observed from

occurring. This figure was reduced to 11% after adjustment for the adult coronary risk factors, and 7% after further adjustment for height.

9.2 Introduction

9.2.1 Background

Social inequalities in the incidence of CHD in the UK have been documented for many years.[53;552] In the United Kingdom, these differences are most often calculated according to occupational measures of social class (e.g. the Registrar General's six category classification).[442] In recent years, though absolute decreases in CHD rates have been observed in categories, the rate of decline has been greater amongst the higher social classes (I, II and IIINM) than in the lower classes, and hence the relative difference between those at the top and those at the bottom of the social scale has increased.[54] However, despite considerable emphasis being placed in recent public health policies on reducing social class inequalities in CHD,[54;55] several important aspects of social class differences remain unresolved. First, in epidemiological studies that relate adult social conditions to subsequent disease risk, adult social class (as determined by job occupation) is often used as a convenient and available indicator of the underlying socioeconomic factors. However, even as an approximate index of socioeconomic status, adult social class is not always precisely categorised and may also change over time. Though these factors will lead to underestimation of the true extent of social class differences in CHD, the extent of such underestimation has not been fully resolved.[553] In addition, the relative contribution of established risk factors to social inequalities in coronary heart disease remains uncertain. It is also possible that novel risk factors including early life influences may be important causes of socioeconomic differences in CHD.[378] However the extent to which these factors may "explain" these risk inequalities, as well as the potential importance of identifying new factors influencing social inequalities, is unclear.

The choice of which measure of association to use when presenting evidence from observational studies can also greatly affect the interpretation and importance of the conclusions. For social inequalities in CHD, many previous studies have used the relative risk to display risk differences between social class groups, and have tended to overlook the more important epidemiological index of the population attributable risk fraction. The

PARF is more relevant to the measurement of risk inequalities in public health situations as it indicates the total impact that control of the risk factor in the population could have on future disease rates. Although social class levels associated with high risks of disease cannot be "eliminated" in the same sense that other coronary risk factors could be (e.g. high blood cholesterol, high blood pressure, cigarette smoking), the assessment of a population attributable risk fraction for social class does allow the estimation of how much CHD could potentially be prevented if the social class divide could be narrowed or eliminated.

9.2.2 Objectives

In the analyses presented in this chapter, the magnitude of social class differences (manual *vs* non–manual occupations) in the incidence of major CHD in men initially free from CHD are estimated, before and after correction for within–person variation in the assessment of occupational social class. The extent that any observed differences can be explained by the established coronary risk factors is calculated, and the population attributable risk fraction for social class (before and after adjustment for the established coronary risk factors) is presented. In a subsidiary analysis, the additional effects of reducing CHD risks to those of social class I are considered, and the effect that within–person variation in the established risk factors has on the estimated contribution of these factors to social differences in CHD are assessed.

9.3 Methods

Similarly to chapters 7 and 8, the analyses presented in this chapter are restricted to men with no baseline evidence of CHD. The primary analyses also exclude men in the Armed Forces. In contrast to chapters 7 and 8 however, major CHD events over the full 20–year period are used in the analyses in this chapter. This is because the repeated information on social class was recorded over a twenty–year interval.

9.3.1 Time to major CHD events by social class status

Kaplan–Meier curves stratified by social class were used to display the differences in major CHD by occupational social class over 20 years. 20–year CHD major event rates were

calculated per 1,000 person years of exposure and directly standardised to the age distribution of the entire cohort. Relative hazards were estimated using Cox proportional hazards regression before and after adjustment for a range of coronary risk factors: blood pressure, blood cholesterol, body mass index, FEV1 and "usual" or "average" exposure to cigarette smoking, physical activity level and alcohol intake (see chapter 5 for definitions of these variables). Adult height was also included in analyses as a marker for adverse socioeconomic conditions during childhood. The proportion of the social class divide "explained by" each risk factor was estimated through the equation $(\beta_0 - \beta_1)/\beta_0$, where β_0 is the age adjusted (log) hazard ratio for manual social class, and β_1 the log hazard ratio after adjustment for the coronary risk factor.[554] Approximate 95% confidence intervals for this statistic were calculated using bias–corrected bootstrap re–sampling of size 1000 to estimate the upper and lower limits.[473]

9.3.2 Error in the ascertainment of occupational social class

Two distinct sources of within–person variation in occupational social class have been considered: (1) misclassification of occupational social class at baseline; and (2) true changes in social class within the first 10 years of follow–up. Throughout this chapter, the combination of these influences is referred to as "social class imprecision". The magnitude of social class imprecision was assessed by comparing the social class measurements taken at baseline with those recorded at the 20–year follow–up (survivors only); this analysis was restricted to subjects who reported at follow–up that they had been in the same occupation for at least 10 years, thus capturing the second component of imprecision mentioned above. [i] The proportion of individuals measured imprecisely (τ) was then estimated through the equation

$$\hat{\tau} = 0.5 - 0.5\sqrt{\frac{N-2n}{N}} \tag{9.1}$$

where N individuals are measured twice and there are n disagreements between the baseline and follow–up measurements[23]. The adjusted log hazard ratio β^* was then obtained

[i] NB/ While men who changed occupation during the second decade of follow–up were excluded when estimating the extent of within–person variation in occupational social class, they were not excluded from the analyses of social class risk–associations presented in this chapter.

from the observed estimate β through the equation

$$\beta^* = \frac{\beta}{1-2\hat{\tau}} \tag{9.2}$$

with approximate confidence intervals calculated from formulae that take into account the variability in both β and $\hat{\tau}$.[24] The effect of within–person variation in other CHD risk factors (particularly total cholesterol and blood pressure) on the estimated relative hazard for social class was assessed using the methods of Rosner *et al.*[26]

9.3.3 The population attributable risk fraction for social class

The population attributable risk fraction for manual social class may be defined as the proportion of all disease events in the population that can be attributed to the excess risks experienced by manual men over those of non–manual men. This fraction is calculated through the equation $p(RR-1)/(1+p(RR-1))$, where p is the proportion of manual men in the population and RR is the relative risk of major CHD for manual men relative to non–manual men. For simplicity, the relative risk of major CHD for manual men relative to non–manual men was approximated by the relative hazard from the Cox proportional hazards model.[555] These estimates were found to be very similar to those obtained when more complex prediction–based approaches (similar to those used in chapter 7) were used to estimate the true relative risk. Approximate 95% prediction intervals for the population attributable risk fraction (before and after correction for measurement imprecision in social class) were calculated using bias–corrected bootstrap re–sampling.

9.4 Results

9.4.1 Risk factors by social class

Occupational social class at baseline was recorded for 6,563 out of the 6,576 men with no baseline evidence of CHD (99.8%). Of these men, 6,386 (97.3%) were not in the Armed Forces, 3,686 of whom (57.7%) were classed as "manual" and 2,700 (42.3%) as "non–manual". The baseline characteristics of the men by occupational social class are shown in Table 9.1. Manual men were, on average, slightly older than non–manual men, had

higher body mass index and blood pressure, were considerably more likely to be current cigarette smokers and heavy drinkers and less likely to be physically active. They were also shorter on average than non–manual men and had poorer lung function (FEV1, height standardised). Manual men did however have lower average total cholesterol levels than non–manual men (6.21 mmol/L vs 6.37 mmol/L; p<0.001).

9.4.2 Imprecision in the assessment of occupational social class

Table 9.2 shows how social class recorded at baseline compares with social class ascertained at 20 years for men without CHD at baseline who, at the 20–year screening, claimed to have worked in the same job for 10 years or more ($N = 3,113$). From these data, the probability that a single baseline assessment of occupational social class (categorised as manual/non–manual) would be the same as the individual's true occupational social class ten year later was estimated to be 0.92. The differences in reported occupational social class between the baseline and follow–up assessments are displayed fully in Figure 9.1. There was a tendency for individuals initially at the lower end of the social spectrum to be "upwardly mobile" (less than 40% of men initially classed as social group V were categorized into this group at follow–up), though it can be seen that, overall, a similar number of men from each social class switched from manual to non–manual by the 20–year assessment, and vice versa. There were no significant differences in age, blood lipids, blood pressure, body mass index, cigarette smoking or physical activity between men whose "manual/non–manual" classification was the same at baseline and follow up ($n = 2,668$) and those men whose classification was different ($n = 445$), or between men who were 'upwardly mobile' ($n = 569$) compared with men who were "downwardly mobile" ($n = 662$). Men living outside the South of England at baseline were more likely to be "upwardly mobile" than men living in the South however (49% *vs* 41%; $p = 0.019$).

9.4.3 Social class variation in major CHD

20–year major CHD event rates by baseline social class

After 20 years of follow up, 580 men in manual occupations (15.7%) and 327 men in non–manual occupations (12.1%) had had a major CHD event. Figure 9.2 displays 20–year major CHD event rates and Kaplan–Meier cumulative incidence curves by baseline occupational social class (in each of the six categories and combined manual *vs* non–

manual). Major CHD event rates increased from 5.8 per 1,000 person years in social class I to 10.2 per 1,000 person years in social class V. For all manual social classes considered together, the age–standardised 20–year major CHD event rate was 9.4 per 1,000 person years; for all non–manual social classes the age–standardised 20–year major CHD event rate was 7.2 per 1,000 person years. These differences in the incidence of major CHD by baseline social class groups can clearly be seen in the corresponding Kaplan–Meier cumulative incidence curves.

Relative hazard of major CHD by social class

The relative hazards of major CHD for men in manual occupations relative to men in non–manual occupations (before and after adjusting for the established coronary risk factors and before and after correcting for imprecision of social class) are shown in Table 9.3. The observed age adjusted hazard ratio for manual men relative to non–manual men before correction for social class imprecision was 1.34, which attenuated to 1.19 after adjustment for the adult coronary risk factors (blood cholesterol, blood pressure, body mass index, cigarette smoking, alcohol intake, physical activity and lung function). Correcting for imprecision of social class increased these estimates to 1.41 (95% CI 1.20 to 1.66) and 1.23 (95% CI 1.03 to 1.46) respectively; indicating a 41% (95% CI 18% to 79%) reduction in the magnitude of the social class differences. Once height was also taken into account, the observed differences between manual and non–manual social classes became non–significant, and 61% (95% CI 30% to 107%) of the original differences between social classes was accounted for. [ii] Of the established adult coronary risk factors, cigarette smoking accounted for the greatest proportion of the CHD differences between social classes, 38% (95% CI 23% to 78%). Physical activity explained 16%, systolic blood pressure 15%, and FEV1 (height standardised) explained 13%, while HDL cholesterol, diastolic blood pressure, body mass index and alcohol intake each explained less than 10% of the observed differences. Adjustment for total cholesterol increased the corrected hazard ratio for social class from 1.41 to 1.53. This was because total cholesterol levels were on average lower for the manual men (see Table 9.1). Thus the proportion of the differential CHD risks by

[ii] NB/ The interpretation of the 107% upper confidence limit is that if these risk factors had been distributed equally across social classes, then true CHD risks in the population of manual social classes may have actually been lower than CHD risks in the population of non–manual social classes (as opposed to the same, i.e. 100% explained, or higher, i.e. <100% explained)

social class explained by total cholesterol was -22% (indicating that the age adjusted log hazard ratio would have been 22% greater had total cholesterol levels been the same in the two groups). Height made a notable additional contribution to the social class gradient, explaining 26% of the difference. This increased the proportion explained by all factors combined from 41% to 61%.

9.4.4 PARF estimates for social class

Table 9.4 shows the population attributable risk fraction for social class (manual vs non–manual) before and after correction for imprecision in social class status and before and after adjustment for the adult coronary risk factors. For major CHD events, the age adjusted PARF for manual social class was 19% (95% CI 11% to 28%) after correction for social class imprecision. If men from manual occupations had had the same average levels of blood cholesterol, blood pressure, body mass index, physical activity, FEV1, alcohol intake and smoking rates as men from non–manual occupations, the "imprecision adjusted" PARF would have been 12% (95% CI 2% to 21%). Further adjustment for height reduced this estimate still further to 8% (95% CI -2% to 18%).

9.4.5 Subsidiary analyses

Effect of redefining the "low–risk" group

When the "low–risk" group was re–defined to include only social class I (rather than all non–manual social classes), the hazard ratios for the high–risk group (social classes II – V) relative to the low–risk group (social class I) and the corresponding PARF estimates increased. The uncorrected age adjusted hazard ratio was 1.51 (95% CI 1.14 to 2.00), which attenuated to 1.26 (95% CI 0.94 to 1.70) after adjustment for the adult risk factors and 1.18 (0.88 to 1.59) after further adjustment for height. The proportion of the difference explained by all adult risk factors and height was almost identical to that observed when comparing manual men with all non–manual men. The uncorrected age adjusted population attributable risk fraction estimate based on this new classification was 32% (95% CI 11% to 48%), which reduced to 19% (95% CI -6% to 39%) after adjustment for the adult risk factors and 14% (95% CI -12% to 35%) after further adjustment for height. No attempt to correct these estimates for within–person variation in occupational social class was made as, in this case, it would not have been reasonable to assume that mis-

classification was non–differential. To correct for within–person variation in this setting would have required repeated information on occupational social class during the study period (for instance, at Q5, Q92 and Q96).

Effects of within–person variation in the established CHD risk factors

The results presented in Table 9.3 show the extent to which the established coronary risk factors "explain" differences in the incidence of major CHD by occupational social class. For cigarette smoking, physical activity and alcohol intake, these analyses were based on exposure measurements that accounted for within–person variation in these factors (i.e. the exposures derived in chapter 5 were used in analyses instead of the baseline exposures). Table 9.5 shows the proportion of social differences explained by these factors when baseline levels were used in analyses. For cigarette smoking, failure to take within–person variation into account would have led to underestimation of the importance of cigarette smoking in explaining CHD differences (it would have been estimated that 33%, rather than 38%, of the social differences could be explained by differences in smoking). However, for physical activity and alcohol intake, the use of "average" exposures in analyses did not increase (or decrease) the estimated importance of these factors.

For the continuous risk factors (blood cholesterol, blood pressure, etc.), the possible effects that within–person variation (regression dilution bias) may have had on the estimated contribution of these factors to social differences in CHD is more difficult to quantify. Using the "expected conditional" values of these measures in analyses instead of the baseline levels (as in chapter 6) would have had no effect on the estimated proportions explained, because though the relative hazards for these factors would increase, if the expected values were calculated across the whole population then the estimated differences in these levels between social classes would decrease, thus cancelling out any effect. However, if the estimated differences in the established risk factors between social classes (shown in Table 9.1) did represent true underlying usual differences, then the individual contribution of the continuous risk factors to social differences in CHD risk would probably have been underestimated. However, when considered together, it is likely that because of the effects of regression dilution bias in blood pressure and total cholesterol would be acting in opposite directions (since men from manual occupations had higher blood pressure but lower total cholesterol), any effects would cancel each other out.

Differential associations by age

In order to assess whether or not the associations between major CHD risk and social class differed according to age, analyses were performed separately for men aged under 50 years at baseline, and men aged 50 years or over. Risk differences between manual and non–manual social classes appeared to be stronger in younger men; the (baseline) age–adjusted relative hazard of major CHD for manual versus non–manual men was 1.62 (95% CI 1.28 to 2.05) in men aged under 50 and 1.21 (95% CI 1.02 to 1.43) in men aged 50 or over. The corresponding PARF values (adjusted for age only) were 26% and 11% respectively. After adjustment for "usual" levels of the adult coronary risk factors these risk associations were reduced by 50% in men under 50 and by 32% in men aged 50 or over to, respectively, 1.27 (0.99 to 1.64) and 1.14 (0.95 to 1.36) – the corresponding PARF values decreased to 17% and 5% respectively. Further adjustment for height reduced the relative hazards to 1.22 (0.94 to 1.58) and 1.05 (0.87 to 1.26) and the PARFs to 11% (in men aged under 50) and 3% (in men aged 50 or over).

9.5 Discussion

9.5.1 Interpretation of findings

Among middle–aged British men with no previous evidence of CHD, the age–adjusted relative hazard of major CHD for men in manual occupations compared with men in non–manual occupations between 1978/80 and 1998/2000 was 1.41 (95% CI 1.20 to 1.66) after correction for imprecision of social class measurement. This was reduced to 1.23 (95% CI 1.03 to 1.46) once differences in the adult coronary risk factors were taken into account. This risk difference appeared to be greater in younger men than in older men. In the BRHS, if all manual men had experienced the same baseline levels of risk as non–manual men, then 19% of all first major CHD events would have been prevented. After taking account of adult coronary risk factors, the population attributable risk fraction for manual social class was 12% for major CHD cases, indicating that if social inequalities in adult coronary risk factors could be eliminated, the remaining risk differences between manual and non–manual men would account for approximately one in nine major CHD cases during middle age. This figure was reduced to 8% after further adjustment for height. The effects of "within–person variation" in the established coronary risk factors on the

estimation of the proportion of social class differences explained by these factors was fairly small, though cigarette smoking did explain a greater proportion of the difference after within–person variation was taken into account (38% *vs* 33%).

9.5.2 Validity of analyses

Using individuals who, at the 20–year screening, claimed to have been working in the same job for at least 10 years, it was possible to correct for both random misclassification of social class status at baseline and true changes in social class during the first ten years, thus enabling estimation of associations between "usual" social class held throughout the study period and disease risk over the study period. Although it is possible that the onset of coronary heart disease during the study may have led to a change in social class for some individuals, separate examination of social class changes in subjects developing CHD during the follow–up period suggested that any such effect would be small. The primary analyses used occupational social class as the most convenient marker for socioeconomic conditions. In a subsidiary analysis, this measure was compared with two other measures (car ownership and housing tenure) both of which were available from the 5–year follow up questionnaire (Q5). Though owning a car and being an owner–occupier were related to lower subsequent CHD risk (between 5 and 20 years), neither was more strongly related than adult social class (over the same period). However, it is recognised that a combined baseline measure of the three socioeconomic factors might have been able to predict CHD outcome better than any single measure in isolation. Social class, though the most widely used, is only one potential measure of socioeconomic status and a combination of different socioeconomic measures may encapsulate different "social class dimensions" better than any single measure in isolation.[556;557] In addition, area–based measures of social deprivation may contribute additional socioeconomic information over and above that obtained from "individual–level" factors,[558] though these types of analyses are beyond the scope of this thesis.

9.5.3 Comparison with other studies

The extent to which social class differences can be explained by established coronary risk factors has been assessed in a number of studies.[362;448–451;559;560] In an earlier report from the British Regional Heart Study in 1987,[448] the rate ratio of CHD mortality for all man-

ual men relative to all non–manual men over 6 years of follow–up was estimated to be 1.44. This was attenuated to 1.24 after adjustment for cigarette smoking, serum total cholesterol, blood pressure and physical activity (a 41% reduction on the log scale). In comparison, a recent 25–year follow–up of men in the Whitehall I study[450] found that cigarette smoking, blood pressure, total cholesterol and glucose together accounted for 56% of the CHD risk difference between low–risk men in the lowest and highest grades of employment. In the Chicago Heart Association Detection Project in Industry study, the relative odds of CHD death within five years for college graduates relative to those not graduating from high school decreased by 33% after adjustment for blood pressure, cholesterol, smoking, BMI and ECG abnormalities;[559] a slightly higher estimate was obtained for 20–year mortality rates of men in the Chicago Peoples Gas and Western Electric Studies.[559] A wider range of risk factors including HDL cholesterol, triglyceride and fibrinogen was found to account for 39% of the CHD difference between social classes in the Scottish Heart Health Study (SHHS),[451] the same estimate as was found in our current analysis (Table 9.3). Further analysis of the SHHS however estimated that 14 risk factors (including the established factors as well as further risk factors including type A personality, fibrinogen and vitamin C consumption) together explained over 70% of the CHD risk differences observed between "renters" and "owner occupiers".[560] Consistent with the BRHS, average total cholesterol levels in the SHHS (as well as in Whitehall I) were greater in men from the higher social classes. Therefore, had total cholesterol levels been the same across the whole population, the CHD differences between social classes would have been greater than was observed (by approximately 22% in the BRHS). However, when assessing the overall combined contribution of the established coronary risk factors to the social class divide, this significant inverse relation with total cholesterol was excluded in the SHHS analyses, whereas in the analysis presented here, it was included. Though the estimated combined contribution of the remaining factors would have been somewhat larger than 41% had total cholesterol been excluded, since the hypothesis was defined *a priori*, it was felt that total cholesterol should remain. Finally, in an ecological study of risk factor prevalence and cause specific mortality in 403 local authority districts in England and Wales in 1992, cigarette smoking was estimated to account for as much as 85% of the observed difference in ischaemic heart disease mortality between the most and least deprived areas.[561] However these findings should be compared with those of this

chapter cautiously because of the ecological fallacy.[562]

When risk factors were considered individually, cigarette smoking was found to account for the largest proportion of observed difference in CHD by social class, explaining 38% of the CHD gradient (see Figure 9.3). Height explained 26% of the variation, and the remaining risk factors each explained 16% or less. The observation that height is the second most important determinant of the social class gradient in CHD suggests that early life factors may be important in the development of social differences in CHD – an observation consistent both with the inverse relationship between height and CHD in individuals[169;362;377;390–394;563] and with recent evidence that childhood socio–economic environment may be directly related to the risk of CHD.[452–454] The addition of height to the adult CHD risk factors resulted in 61% of the CHD difference being explained. Some of the remaining association may be explained by additional contributions made by homocysteine level,[300;301;506] by job stress[361] or by other unmeasured early life factors.[378]

9.5.4 Effect of correction for within–person variation in major risk factors

Estimating the long–term social inequalities remaining after adjustment for baseline risk factors depends on the assumption that the effects of within–person variation do not differ by social class. This is unlikely to be true for cigarette smoking because of differential rates of cigarette cessation.[564] In the British Regional Heart Study, 33% of non–manual smokers had given up by five years compared with 24% of manual smokers. The effect of these (and subsequent) differential changes in smoking exposure on the extent that smoking "explains" social class differences in CHD was to increase the estimate from 33% to 38% (see Table 9.5). While taking account of within–person variation in physical activity and alcohol consumption had virtually no effect on the proportion explained by these factors (Table 9.5), it is likely that differential changes in mean total cholesterol by social class during the study may also have led to *underestimation* of the extent that social inequalities are truly explained by the established risk factors. This is because while mean total cholesterol was *higher* in non–manual social classes than manual social classes in 1978–80, little to no differences are observed in contemporary adults (Health Survey for England, 1998).[565] Therefore, the *average* difference in total cholesterol between manual and non–manual men in the BRHS during the period 1978/80 to 1998/2000 is likely to

have been lower than that estimated using baseline data. [iii] Therefore, the (negative) effect of adjustment for total cholesterol in the analyses presented in Table 9.3 is likely to have had a greater effect than it would have had any differential changes in total cholesterol by social class been measured and taken into account, and hence the proportion of social class differences explained all major risk factors would have been greater. Similarly, other differential changes in risk factors by social class would most likely lead to underestimation of the extent that they "explain" social inequalities in CHD (because they are more likely to result in improvements in the risk profile of non–manual men relative to manual men).

9.5.5 Conclusions: social class inequalities

Social class inequalities in CHD risk are underestimated when based on measurements taken at a single point in time, irrespective of which marker of social class is used to represent the underlying socioeconomic factors. However, even taking account of measurement imprecision, the contribution of occupational social class to overall CHD risk is likely to be relatively modest. The analyses in this chapter suggest that approximately one–fifth of all major CHD events in the BRHS would have been prevented or postponed if the average risks of non–manual social classes had been experienced by the whole population. A reduction in CHD rates of this magnitude would be both important and desirable. However, compared with the reduction in CHD risk which could be achieved by population–wide changes in just two of the three most important causal factors for CHD (blood cholesterol and blood pressure; refer to chapter 8), this is a fairly limited reduction. Population–wide strategies to reduce major CHD risk factors across all social class groups are therefore likely to produce greater benefits for CHD prevention than strategies designed specifically to reduce social inequalities in CHD. However, to secure both effectiveness and equity in CHD prevention (and in the reduction of all–cause mortality), these population wide measures would logically include specific measures to encourage smoking cessation and physical activity among socially disadvantaged groups.

[iii] this is indirectly supported by the observation that at the 20–year rescreening of surviving BRHS participants, mean total cholesterol was 6.0 mmol/l in *both* manual and non–manual men

Table 9.1: Mean baseline characteristics of the 6,563 men with no evidence of CHD at baseline, stratified by (baseline) social class group.

Risk factor	Registrar General's six category classification								Non–manual/Manual	
	I (n=530)	II (n=1532)	IIINM (n=638)	IIIM (n=2763)	IV (n=667)	V (n=256)	Army (n=177)	P–value*	I,II or IIINM	IIIM, IV or V
Age (years)	49.0	49.7	50.1	50.1	50.2	49.8	49.7	<0.001	49.6	50.1
Total cholesterol (mmol/L)	6.34	6.38	6.36	6.22	6.19	6.16	6.46	<0.001	6.37	6.21
HDL cholesterol (mmol/L) †	1.15	1.13	1.11	1.12	1.13	1.14	1.09	0.12	1.13	1.12
BMI (kg/m^2)	24.9	25.3	25.4	25.6	25.3	25.3	25.7	0.001	25.3	25.6
SBP (mm Hg)	141	143	146	147	146	148	140	<0.001	143.1	146.7
DBP (mm Hg)	80.7	81.1	83.1	82.8	83.0	82.7	80.3	<0.001	81.5	82.8
Height (cms)	176	175	174	172	171	171	174	<0.001	175	172
FEV1 (ml/sec) ‡	358	348	338	330	318	323	332	<0.001	347	327
Current smokers (%)	19.7	30.4	35.8	46.0	50.8	55.5	62.1	<0.001	29.6	47.5
Moderately active (%)§	52.6	47.7	41.5	34.4	30.9	26.2	35.5	<0.001	47.2	33.2
Heavy drinkers (%)§	1.5	3.2	2.8	4.2	4.5	3.9	5.1	<0.001	2.8	4.2

* P–value for test of trend across six categories (excluding Army); †Geometric mean; ‡height standardised; §After taking within–person variation into account.

Table 9.2: Repeat determination of social class status over a 20–year period for men without evidence of coronary heart disease at baseline who, at the 20-year screening, had worked in the same job for at least 10 years ($N = 3,113$)

	20-year follow–up		
Baseline social class	Manual	Non-manual	Total
Manual	1305	257	1562
Non-manual	188	1363	1551
Total	1493	1620	3113

Table 9.3: Relative hazard of major CHD over 20 years (manual vs non–manual social class) before and after taking misclassification of social class into account. All models are adjusted for age

		Uncorrected	Corrected
Variables adjusted for	Percentage explained *	HR (95% CI)	HR (95% CI)
None	-	1.34 (1.17,1.54)	1.41 (1.20,1.66)
Total cholesterol	-22%	1.43 (1.25,1.64)	1.53 (1.30,1.79)
HDL cholesterol	5%	1.32 (1.15,1.52)	1.39 (1.18,1.63)
Systolic blood pressure	15%	1.28 (1.12,1.47)	1.34 (1.14,1.58)
Diastolic blood pressure	9%	1.31 (1.14,1.50)	1.37 (1.17,1.61)
Cigarette smoking §	38%	1.20 (1.05,1.38)	1.24 (1.05,1.45)
Body mass index	7%	1.31 (1.15,1.51)	1.38 (1.18,1.62)
Physical activity §	16%	1.28 (1.11,1.47)	1.33 (1.13,1.57)
Alcohol §	6%	1.32 (1.15,1.51)	1.39 (1.18,1.63)
FEV1	13%	1.29 (1.12,1.48)	1.35 (1.15,1.59)
Height	26%	1.24 (1.08,1.43)	1.29 (1.10,1.52)
Smoking, blood pressure & blood cholesterol	34%	1.21 (1.05,1.40)	1.26 (1.06,1.48)
Adult coronary risk factors¶	41%	1.19 (1.03,1.38)	1.23 (1.03,1.46)
Adult coronary risk factors plus height	61%	1.12 (0.96,1.30)	1.14 (0.96,1.36)

* Percent reduction in the log hazard ratio relative to the model that adjusts only for age. Correction for imprecision in social class measurement has no effect no these estimates; §Usual exposure levels over the follow–up period; ¶Cigarette smoking, blood pressure, blood cholesterol, body mass index, physical activity, alcohol and height standardised FEV1; HR = hazard ratio; CI = confidence interval.

Table 9.4: Population attributable risk fraction for manual social class for men with no baseline evidence of CHD. Estimates are presented before and after correction for measurement imprecision in social class status. All estimates are adjusted for age.

Factor included	Uncorrected		Corrected	
	PARF	95% CI	PARF	95% CI
None	16%	(10%,24%)	19%	(11%,28%)
Adult coronary risk factors*	10%	(2%,18%)	12%	(2%,21%)
Adult coronary risk factors plus height	6%	(-2%,15%)	8%	(-2%,18%)

* cigarette smoking, blood pressure, blood cholesterol, body mass index, physical activity, alcohol and height standardised FEV1.

Table 9.5: Percentage of social class differences explained by the established coronary risk factors before and after correction for within–person variation in the established CHD factors.

Factor included	Before correction	After correction*
Cigarette smoking	33%	38%
Physical activity	15%	16%
Alcohol intake	6%	6%
All coronary risk factors §	39%	41%

* these figures are the same as those shown in Table 9.3; §cigarette smoking, blood pressure, blood cholesterol, body mass index, physical activity, alcohol and height standardised FEV1.

Figure 9.1: Movement in social class amongst 3,113 men with no baseline evidence of CHD who (at 20 years) claimed to have been working in the same job for at least 10 years

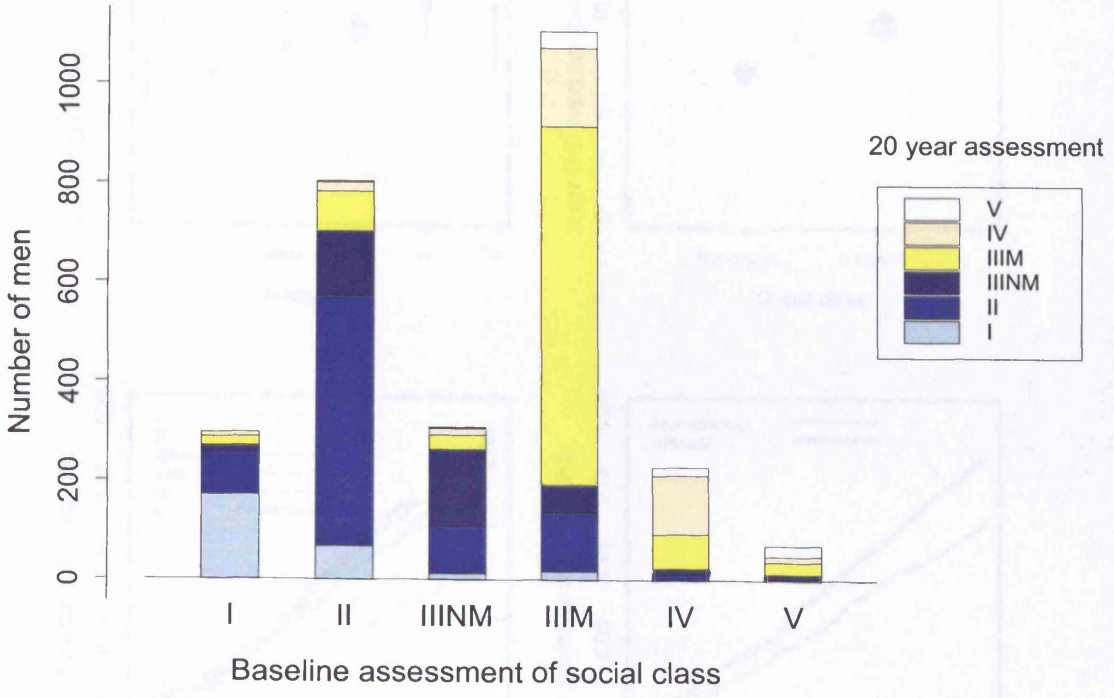


Figure 9.2: Baseline social class differences in major CHD over 20 years among 6,386 men not in the Armed Forces and with no baseline evidence of CHD

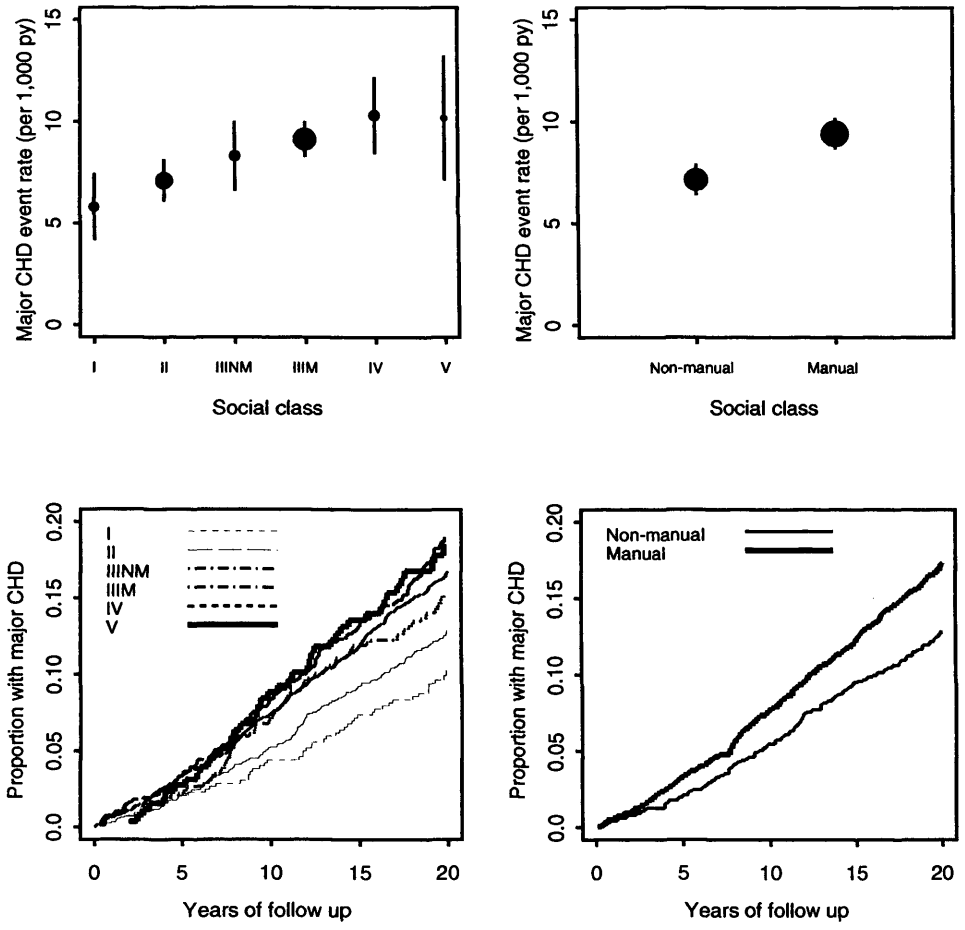
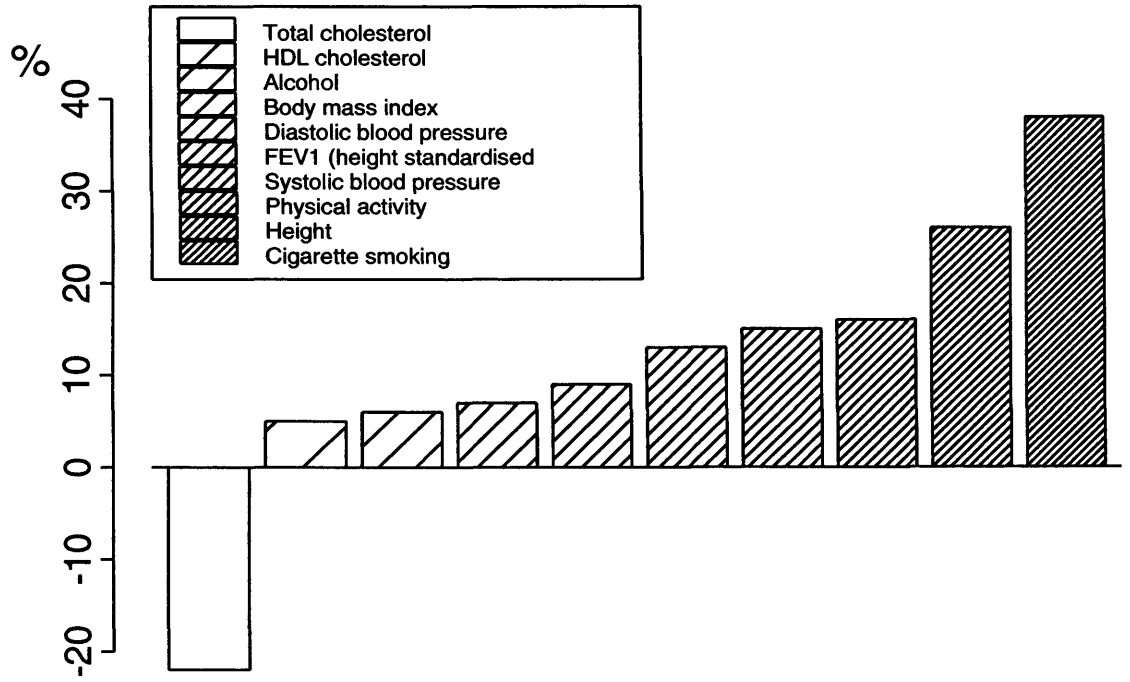


Figure 9.3: Percentage of the difference in CHD risk between manual and non–manual men that can be explained by the established CHD risk factors



Chapter 10

Implications of the findings

10.1 Summary

The findings of this thesis have implications for the design and analysis of epidemiological studies, and for the prevention of coronary heart disease. The thesis results demonstrate the extent of within–person variation in coronary risk factors typically observed over the duration of a cardiovascular epidemiological study, and the effects that this variation can have on estimated relationships between risk factors and disease, both in a single and a multiple covariate setting. They also highlight the difficulties of assessing the effectiveness of CHD prevention policies, when estimates are based on single risk factor measurements taken at a baseline assessment, and the relative importance that should be attributed to reducing inequalities in CHD. The most important implications from this thesis are: (i) that epidemiological studies should recognise and, where appropriate, plan to correct for the effect that within–person variation has on estimated disease relationships; (ii) that the three strongest risk factors for CHD account for substantially more than half of CHD cases in British men; (iii) that currently recommended strategies for the primary prevention of CHD are likely to have a very limited effect on population levels of CHD unless used much more widely; (iv) that even small reductions in key causal risk factors throughout the population could have a large effect on reducing CHD; and (v) that social inequalities in CHD are larger than previously estimated, though reduction of social class inequalities is unlikely to provide a means through which CHD could be substantially reduced, even if they could be entirely eliminated.

10.2 Introduction

This chapter considers the implications of the thesis results, first for epidemiological studies (section 10.3) and second for public health (section 10.4). The chapter builds on the results, discussion and conclusions of previous chapters, although in some sections additional relevant literature is introduced.

10.3 Implications for epidemiological studies

In this section, the implications of the results of this thesis for the design and analysis of future epidemiological studies are presented. Particular emphasis is paid to: (1) identifying relationships between long–term risk factors levels and disease risk for single continuous risk factors; (2) identifying relationships between average levels of categorical risk factors and disease risk; and (3) estimating the simultaneous contribution of several risk factors to disease risk while taking account of within–person variation in more than one risk factor.

10.3.1 Estimating disease relationships for continuous risk factors

The effect that within–person variation in a single risk factor has on its estimated relationship with disease risk is most easily described for the continuous risk factor that displays a log–linear "dose–response" relationship with the risk of disease. These are risk factors for which a unit increase in exposure to the factor leads to a constant proportional increase (or decrease) in risk, irrespective of the starting level. For coronary heart disease, two of the most important risk factors, blood cholesterol and blood pressure, display such relationships.[70;113] CHD risk increases linearly with increasing blood pressure, serum total cholesterol, and the ratio of total to HDL cholesterol, and decreases linearly with HDL cholesterol (Chapter 6). Under these circumstances, the estimated relations between "baseline levels" of these risk factors and incident CHD underestimate the true strength of relation between average levels and disease risk, because of regression dilution bias (Chapter 3). The extent of this bias is determined by the degree that observed differences in baseline levels truly reflect long–term differences between individuals. Therefore, this bias may be corrected for if an estimate of the true "between–person" variation to the apparent variation (the regression dilution ratio) can be obtained. Reassuringly, correction for regression dilution bias in a single continuous risk factor does not alter the statistical

significance of its estimated relationship with disease risk. The reason for this is that, in its simplest form, the correction method simply multiplies the observed association by a correction factor. Therefore if the confidence interval for the observed association contains zero (consistent with the null hypothesis) then so will the confidence interval for the corrected association.

Estimates of the regression dilution ratio (RDR)

Estimates of the extent of regression dilution bias are important if unbiased estimates of associations between usual risk factors and disease outcomes are to be obtained. Long–term prospective studies therefore need to recognise these effects, and adjust for them using appropriate estimates of the regression dilution ratio. This can be done in a variety of ways (see chapter 3), but each requires the risk factor to be remeasured after a particular duration of follow–up. However, before correction for regression dilution bias is applied, several important issues need to be considered.

1. **Over what follow–up interval should the RDR correction factor be evaluated?**

   The decision on when to re–measure individuals depends on the length of follow–up over which risk–associations are to be calculated. The approach taken in this thesis has been to relate disease risk over a particular follow–up period to usual risk factor levels around the mid–point of that interval.[19] Therefore, for the 20–year associations, major CHD risk was related to usual risk factor levels after 10 years of follow–up, while for the 10–year associations, risk was related to usual exposure levels after approximately 5 years of follow–up. Variations of this approach have been used elsewhere.[113;533] In the Prospective Studies Collaboration, the risk of CHD death in each 10–year age group was related to the estimated usual blood pressure level at the start of that 10–year period,[113] while in a recent analysis of the Whitehall I study, CHD mortality was related to usual blood pressure and blood cholesterol levels at about 5 years before death.[533] By taking into account the observation that individuals who die from CHD later in life have, on average, been followed for longer periods of time, the authors of these reports were able to perform "time–dependent" correction for regression dilution bias, applying different correction factors for subjects "at risk" during different decades of age. Similarly, by using separate correction

factors for men at risk of major CHD in the first and second decades of follow–up in the BRHS (see §6.4.7), time–dependent correction for regression dilution bias was performed in this thesis. Taking into account the longer follow–up periods for men at risk of a first major CHD event between 1988/90 and 1998/2000, the true relationship between usual blood lipid and blood pressure levels and major CHD risk, rather than diminishing over time as first it seemed, was shown to be fairly stable (see Table 6.4). The reason why it is important to choose an appropriate period over which to correct for regression dilution bias, and the reason why its choice can greatly affect the results, is that for most risk exposures the regression dilution ratio tends to decrease (i.e. the regression dilution effect gets stronger) as the interval over which it is estimated increases (as can clearly be seen in Figures 5.2 to 5.5). If real data are not observed over this "ideal" period then it may still be possible to estimate the likely effects providing that sufficient repeated data over other intervals are available (for instance, the 10–year regression dilution ratio estimates used in chapter 6 were estimated, not from real data over a 10–year period, but from estimates of the regression dilution ratio over periods of 1 week, 4 years, 16 years and 20 years; see Table 6.1).

2. **Should the RDR be estimated from external data or from the same study?**

In this thesis, estimates of the regression dilution ratio over 4, 16 and 20 years were taken from all BRHS men with repeated risk factor measurements available over those periods. The advantage of using these data to correct for regression dilution bias is that they should provide appropriate adjusted point estimates (because they are drawn from the population being studied). The disadvantage is that (unless they are excluded from other analyses) the observations are not independent of the studied population (because they form a subset of the main study). While this doesn't affect the point estimates of the corrected disease–associations, it would mean that confidence intervals for these estimates would be too narrow if independence were to be assumed. This problem can be overcome by using bootstrap re–sampling to calculate the confidence intervals, as this method allows for any correlation between the uncorrected disease–association and the correction factor when calculating confidence intervals for the corrected disease–association. Alternatively, "external" data may be used to calculate correction factors. This may be data from the orig-

inal study participants subsequently excluded from the estimation of uncorrected disease–associations, or else may be from an entirely separate study (e.g. the one–week repeated data in this thesis was taken from a study of men in Islington, North London, who were not BRHS participants). The advantage of using either of these two sources of external data to estimate correction factors is that the estimates would be independent of the uncorrected disease–outcome associations. However, one of two other problems may be introduced depending on which source of external data were used. If a subset of the main study were used, the subsequent exclusion of these participants from the estimation of disease–outcome associations could potentially introduce certain selection biases, which could result in biased estimates. However, if data external to the main study were used, one would need to be sure that the estimated correction factors were truly generalisable to the population under study. In practice, the decision on which data to use to estimate correction factors (and, if from the main study, whether to subsequently exclude it when estimating disease–outcome associations) may depend on the statistical methods available to the researchers, as well as the relative weight that is given to avoiding selection biases on the one hand, and ensuring that the correction factor is appropriate on the other. The latter concern is now addessed in more detail.

3. **Are RDR estimates similar between different populations?**

Ideally, studies that wish to correct for regression dilution bias would probably wish to use repeated risk factors measurements from their own study participants, or at least a subset of the study participants, in order to ensure appropriate estimates of the regression dilution ratio. However, if repeated measurements of individuals are not available (or remeasurement of participants isn't feasible), then estimates from other studies such as the BRHS may be helpful. The estimates of the RDR presented in this thesis were found to be fairly consistent with those from many other studies in many different populations (see Chapter 5; Figure 5.12). Therefore, it may be reasonable to assume that they can appropriately be used to correct for regression dilution bias in other study populations. Furthermore, the 20–year estimates of the regression dilution ratio for the blood lipid and blood pressure indices were found to be reasonably independent of age, social class and town of residence (see Chapter 5; Figures 5.6 to 5.11), providing further support for their likely validity

in other settings. In addition, other studies have found that gender also seems to have little effect on the size of the regression dilution ratio,[19] indicating that the BRHS estimates shown in Chapter 5 may even be valid for producing corrected regression coefficients in studies of women. However, the level of pre–existing CHD in the population may affect the validity of performing such corrections. Consistent with previous studies,[524–526] regression dilution effects were generally found to be greater (i.e. RDR estimates were lower) in men with evidence of CHD at baseline, so that applying correction factors derived from all BRHS men to populations initially without CHD may result in "over–correction" of regression coefficients, i.e. the corrected regression coefficients would be too large. It was for this reason that in chapters 7 and 8 of this thesis, four–year regression dilution ratio estimates were calculated only from men initially without CHD.

4. **Can RDR estimates be applied to the study of other diseases?**

   The concept of regression dilution bias and the methods available to correct for it are quite general. Therefore, providing that interest lies in the relationship between usual risk factor levels over a period of time and disease risk, the regression dilution ratio estimates presented in this thesis do not need to be restricted to the original outcome of interest, in this case major CHD. They would be equally valid for assessing, say, long–term associations with the incidence of stroke, cancer or all–cause mortality, or indeed any outcome sharing at least one of the risk factors presented in this thesis.

10.3.2 Estimating disease relationships for categorical risk factors

Continuous risk factors that display dose–response relationships with disease risk are easily adjusted for regression dilution bias, as their slopes are simply modified by a single correction factor (chapter 6). Similarly, categorical risk factors that can take one of two values can be corrected for misclassification bias through a simple adjustment (chapter 9), provided that the misclassification is non–differential. However, for categorical risk factors that have more than two levels, including risk factors that are grouped into categories for analyses because they display non–linear relationships with disease risk (e.g. alcohol consumption and CHD), and risk factors whose underlying continuous scale is difficult to quantify (e.g. physical activity), the effects of within–person variation on their

relationships with disease risk cannot necessarily be predicted. Therefore, the only way to establish the effect that within–person variation has on disease relationships is to obtain follow–up information on surviving individuals at regular periods throughout the study and to use the information by: (a) fitting changes in risk factors during the study as time–updated covariates; or (b) using the data to recalculate "average exposure" to the factor over a certain period of time. In this thesis, the second approach was used for cigarette smoking, physical activity and alcohol intake in order to derive average exposures to these factors over the 20–year study period (chapter 5). The first approach usually relies on precise information on when changes in risk factors occur and is often used for risk factor changes that are known without measurement error (for instance, the time to CHD death may be greatly influenced by a revascularisation procedure, which could then be fitted as a "known" time–updated covariate). Irrespective of how the follow–up data are used however, if the aim is to identify relationships between risk exposures and "first" incident disease events, then it is important that only follow–up information obtained while the subject remains "event free" should be used in analyses. For instance, a physically inactive man who experiences a non–fatal myocardial infarction after 1 year, and then subsequently becomes active and remains so for the next 19 years would, in the analyses presented in this thesis, be classed as "inactive" in the analysis of first major CHD events.

10.3.3 Identifying "dose–response" relationships

A very important function of an epidemiological study is to try to identify causal risk factors for a particular condition. Though causality cannot be established from epidemiological studies alone, this likelihood would be increased if a risk factor was shown to display a "dose–response" relationship with a disease, meaning that the risk of disease increased (or decreased) progressively with increasing exposure to the risk factor (one of the key criteria for causality proposed by Sir Austin Bradford–Hill).[566] How does within–person variation therefore affect the ability to identify such relationships? For single continuous risk factors, any true dose–response relationships should still be identified when baseline measures are used in analyses (though their strength of relationship would of course be underestimated). However for categorical risk factors this is not necessarily true, and will depend on the nature of the within–person variation ("misclassification"). In the BRHS, the strong "dose–response" relationship between cigarette smoking and CHD was only

observed across the entire spectrum of cigarette smoking exposure after adjustment for within–person variation was carried out (chapter 6). When baseline measures were used in analyses, the increasing CHD risks observed with increasing exposure to cigarette smoking were weaker and were only observed up to a level of 21–39 cigarettes a day; men who smoked at least 40 cigarettes a day had similar observed CHD risks to men who smoked only 1–20 cigarettes a day. Furthermore, though the baseline association between alcohol and CHD indicated moderate benefits from light levels of drinking, and no excess risks from heavy drinking, the interpretation based on average alcohol drinking patterns was quite different, with large excess risks caused by heavy drinking being identified and only a small benefit from light drinking. In both these cases, many of the problems in ascertainment of the true underlying risks from continued exposure to these factors were caused by a high degree of *selective* misclassification of certain levels of exposure, in particular the proportion of men truly exposed to heavy smoking and drinking. Of course, for cigarette smoking and CHD, the large differences in risk between the never–smokers and all current smokers provides the most compelling evidence for causality. However, not all epidemiological studies examine relationships as strong as the cigarette smoking–CHD relationship, and the potential for even large studies to fail to identify a true "dose–response" relationship, or even to incorrectly identify the nature of a true relationship (as appears to be the case for alcohol and CHD, in the BRHS at least), is concerning. Wherever possible, this potential error therefore needs to be taken into account by study investigators, by obtaining follow–up information on individuals and examining the potential effects that risk factor misclassification over time could have on estimated disease relationships.

10.3.4 Estimating "multivariate" relationships

The simple methods of correcting for within–person variation in a single continuous risk factor do not necessarily translate to the case of simultaneous adjustment for several risk factors, when more than one of them is subject to within–person variation.[25–27;521;567] In particular, observed baseline associations may under– or overestimate true risk–associations, even for risk factors not subject to within–person variation. The extent and direction of this bias depends on the correlation between the risk factors. In the analyses considered in this thesis, systolic blood pressure and total cholesterol were found to be fairly independent of one another and, as a result, the simultaneous relations between these factors

and CHD risk after multivariate correction for regression dilution bias were found to be similar to the associations after single adjustment methods were performed (though some small differences were observed, including a slightly reduced estimate of the relative hazards for age and cigarette smoking; chapter 7). Studies that wish to employ univariate correction methods in multiple covariate analyses should therefore do so with caution. The only situation where it can reliably be assumed that this method leads to exactly the same estimates as if multivariate correction techniques were employed is where there is no correlation whatsoever between the explanatory variables (which, in practice, is rarely the case).

10.4 Implications for public health

10.4.1 Implications for the aetiology of CHD

The findings from chapter 6 of this thesis show that for continuous risk factors that display "dose–response" relationships with CHD, failure to correct risk–associations for within–person variation can lead to marked underestimation of their importance. The findings are of particular importance for the aetiology of CHD, and particularly for those established risk factors known to be causally related to CHD, especially blood cholesterol and blood pressure. Before correction for regression dilution bias, the 20–year age adjusted relative hazard of major CHD per unit (1 mmol/l) increase in serum total cholesterol was 1.39 (95% CI 1.32 to 1.46), and per 20 mmHg increase in systolic blood pressure the relative hazard was 1.25 (95% CI 1.19 to 1.32). After correction for regression dilution bias, these estimates increased to 1.63 (95% CI 1.52 to 1.76) and 1.48 (95% CI 1.36 to 1.61) respectively. For systolic blood pressure, this risk difference increased further when restricted to men with no previous evidence of CHD (HR = 1.78; see Chapter 7, Figure 7.2). For the categorical "lifestyle" risk factors, where correction for within–person variation was performed by using information from follow–up questionnaires, the effects of within–person variation were less predictable. For cigarette smoking, approximately a twofold difference in CHD risk was observed between current smokers and never smokers. However, when baseline measures were used in analyses, the "dose response" relationship observed up to 21–39 cigarettes a day was not continued to levels of 40 or more cigarettes a day (see Chapter 6, Figure 6.7). After information from the follow–up questionnaires was taken into account

however, and individuals were reclassified based on "average" exposure to tobacco during the study period, the risk gradient increased and a clear positive dose–response relationship was observed across all levels of cigarette smoking exposure (with men who, on average, truly smoked 40 or more cigarettes a day having over 4 times the risk of major CHD than never–smokers). The benefits of regular, moderate, levels of physical activity were also underestimated when baseline estimates were used in analyses. Compared with inactive men, major CHD risk in moderately active men was 53% lower before, and 68% lower after adjustment for within–person variation. Vigorously active men had half the risk of CHD than inactive men. For alcohol intake, there was a clear limitation of the use of baseline risk factors levels, namely the inability to detect true increased risks associated with continued heavy drinking. This was due to the observation that, of the men who were defined as heavy drinkers at baseline, only approximately one quarter of them were truly heavy drinkers throughout the study period. Comparing major CHD rates between these men and the men who were truly non–drinkers, a 75% excess risk was observed (95% CI 31% to 133%), compared with a 16% lower risk (95% CI 35% lower to 8% higher) estimated from baseline data. The estimated benefits from light to moderate levels of drinking were also reduced when within–person variation was taken into account. These particular findings suggest that even more caution may be needed before any recommendation of alcohol consumption for health reasons is made.

Combined contribution of the three major established risk factors

The overall combined contribution of blood cholesterol, blood pressure and cigarette smoking to premature (before the age of 70) major CHD risk in apparently "healthy" men was examined in chapter 7. Serum total cholesterol and systolic blood pressure were selected as the most informative indices of blood cholesterol and blood pressure respectively, and the combined population attributable risk fraction corresponding to "high" cholesterol, "high" blood pressure and cigarette smoking was calculated for a range of cut–off criteria used to define the high–risk group. The results suggested that effective early primary prevention of CHD focussed towards the control of blood cholesterol, blood pressure and cigarette smoking at the population level could virtually eliminate the disease. Claims that at least half of CHD risk cannot be explained by the established risk factors[30–39] are completely untenable. Had all men had the blood cholesterol and blood pressure levels

of those in the bottom fifth of the distributions, and been non–smokers, then four–fifths of early CHD would have been prevented. Even larger estimates were obtained when the risks associated with previous active smoking or heavy passive smoking were taken into account. Furthermore, these estimates do not take into account control of other closely–related risk factors for CHD, particularly obesity and physical inactivity, which would result in a further increases in the proportion of all CHD cases prevented.

How widely do these findings apply?

Although the results in this thesis apply only directly to men in one reasonably wealthy country with a high prevalence of risk factors and CHD rates which were high by international standards over the follow–up period studied,[568] they are likely to be widely applicable. In particular, it is likely that they apply to women, in whom similar population risk factor distributions, relative risks and population attributable risk fractions apply for blood cholesterol and blood pressure.[42;44] They are also likely to be of considerable relevance to the epidemic of coronary heart disease now emerging in the less affluent countries of the world, in which rising levels of blood cholesterol and blood pressure, and marked increases in the prevalence of cigarette smoking and obesity are prominent.[2] Controlling blood cholesterol, blood pressure and cigarette smoking throughout these populations is needed and should eventually lead to substantial decreases in CHD in the population.

10.4.2 Implications for novel CHD risk factors

The widespread popularity of the "only 50% claim" has helped fuel the notion that other crucially important CHD risk factors remain undiscovered. By 1985, nearly 300 risk factors for coronary heart disease had been identified, and though it is unlikely that more than a small fraction of these actually increase the risk of the disease,[569;570] the possibility that novel risk factors are important causal determinants of CHD risk has gained much support. In particular, there has been considerable interest in the role that nutritional factors (e.g. homocysteine, vitamin C, vitamin E) and chronic infections (e.g. Chlamydia pneumoniae, Helicobacter pylori or specific viral infections) may have, as well as the importance of inflammatory, genetic, social and haemostatic factors (see chapter 2).[36;112;334;344;361;571] There has also been considerable interest in the possibility that fetal nutrition influences CHD risk, possibly interacting with later obesity.[572;573] Relationships between these fac-

tors and CHD risk are also likely to be underestimated because of within–person variation (the extent of this underestimation for the novel risk factors measured in the BRHS is shown in Chapter 5; Figures 5.4 and 5.5), however this underestimation is only important when the association is confirmed to be causal. In addition, though any or all of these factors could be playing a role in the causation of CHD (the evidence for the involvement of homocysteine is particularly compelling),[300;301] such factors would also need to be widely distributed and strongly related to CHD risk to make a substantial independent contribution to the CHD epidemic. Even if this were not the case however, novel risk factors could still be important if they were found to have a direct effect on the major established factors, particularly blood lipids and blood pressure. Slow fetal growth, for instance, has been proposed to influence CHD risk, at least partially, through possible adverse effects on adult blood pressure[132–135] and blood cholesterol,[135;384–386] though two recent systematic reviews have suggested that is not the case for either risk factor.[136;389] Similarly, it is unlikely that blood cholesterol or blood pressure may mediate any relationships between chronic infections and CHD risk.[340;574]

10.4.3 Implications for the primary prevention of CHD

High–risk strategies to prevention

The findings from chapter 8 of this thesis indicate that high–risk strategies based on single risk factor management (of blood cholesterol or blood pressure) would have only a marginal impact on the occurrence of CHD in the population, preventing at best one in ten CHD events from occurring (even if the top fifth of the distribution were all treated). This is consistent with a recent report from the BUPA cohort study, which found that persons in the top 20% of the distribution of systolic blood pressure experienced only 39% of the subsequent coronary deaths that occurred (of which perhaps one fifth may have been prevented or postponed through blood pressure lowering drugs).[575] As was observed in a recent review of interventions to lower blood pressure and cholesterol and their effects on CVD risk,[435] it was observed that by taking multiple risk factors into account, predicted Framingham risks generally provide a more effective measure on which to base treatment decisions than single measures of cholesterol or blood pressure. However, even if multiple risk reducing treatments (with aspirin, a statin, a β–blocker/diuretic and an ACE inhibitor) were provided to individuals at the threshold level currently recommended in

the United Kingdom (10–year risk of at least 30%),[408] only approximately one in nine first major CHD events during middle–age would be prevented. In order to have a large effect on population levels of CHD, multiple risk factor modification would need to be widely used (i.e. multiple combined drugs would need to be prescribed at lower treatment thresholds). This is the same rationale as that underlying the "polypill" approach to prevention, where it has been suggested that everyone in the United Kingdom aged over 55 should be treated with multiple blood pressure lowering drugs, a statin, aspirin and folic acid.[425] In the analyses presented in this thesis, reducing the treatment threshold from 30% to a predicted ten–year CHD risk of at least 20% (as was, until recently, recommended in European guidelines)[4] increased the proportion of the BRHS men defined as "high–risk" to one quarter, and increased the proportion of all premature major CHD events potentially prevented to one third. Further reduction of this treatment threshold to a 15% ten–year risk may have prevented half of all events from occurring, at the expense of treating half the healthy population with multiple combined drugs. Such widespread use of drugs in the population should be advocated with caution however. Concerns regarding the use of drugs as a substitute for a healthy lifestyle, the speculative nature of the assumed effects the drugs would have in primary prevention, the failure to tackle other important risk factors including cigarette smoking, obesity, diabetes and physical inactivity, and the prohibitive costs of such a policy should not be forgotten.[576]

Population approaches to prevention

In comparison to the fairly limited effectiveness of the high–risk approach, small population–wide reductions in total cholesterol and blood pressure would be likely to have a large long–term effect on population levels of CHD. Reducing everyone's blood cholesterol by 0.3 mmol/l and everyone's systolic blood pressure by 7 mmHg (corresponding to a 5% reduction in population mean levels) would be expected to reduce major CHD events by 24%, though the true figure would most likely be greater than this due to additional favourable effects on other risk factors. Reducing both blood cholesterol and systolic pressure by 0.6 mmol/l and 15 mmHg respectively would have led to at least 42% fewer premature major CHD events. Importantly, these sized reductions in cholesterol and blood pressure are fairly small in comparison with the size of the international differences in cholesterol and blood pressure that exist,[545] and should be achievable through concerted population–wide

changes in diet, specifically reducing salt intake and the proportion of calories derived from saturated fat.[104;126;434] However, while many countries have indeed shown steady downwards trends in mean total cholesterol and systolic blood pressure levels over the last 20 years (as demonstrated by the WHO MONICA study of 38 different populations from 21 countries),[577] cigarette smoking rates in women, as well as increasing mean body mass index (particularly in men) have also been observed. Such adverse population–wide changes in coronary risk factors are likely to have significant public health implications in the future.

10.4.4 Consistency with trends in CHD

Since the late 1970's, CHD mortality in the United Kingdom has declined sharply (by approximately 50%),[3] while in other Western countries even larger reductions have been observed (see Chapter 1; Figure 1.2). Though these changes in CHD mortality are due in part to changes in survival following a CHD event (caused by the improvements in coronary care and secondary prevention observed during the 1980's and 1990's),[578] it is the changing coronary event rate (rather than the case fatality rate) that is likely to be the driving force behind these reductions in CHD mortality.[6;579] To what degree, therefore, are changes in risk factors responsible for these reductions? Furthermore, if risk factor changes are responsible, which factors have had the greatest effect and is this consistent with the findings from this thesis? Numerous authors have attempted to estimate the extent that national reductions in CHD mortality rates can be attributed to population–wide changes in major risk factors, and the extent that they may be due to other factors operating at the population level including improvements in the prevention and treatment of CHD.[544;548;578;580–588] In Scotland, for instance, risk factor changes between 1975 and 1994 were estimated to explain approximately half of the decrease in CHD mortality observed over that period.[548] This was predominantly caused by reductions in cigarette smoking between 1975 and 1994, with additional effects attributed to the relatively small secular reductions in cholesterol and blood pressure. Similarly, in New Zealand between 1982 and 1994,[580] in the Netherlands between 1978 and 1985,[584] and in the United States between 1968 and 1976,[585] and between 1980 and 1990,[586] reductions in coronary risk factors were estimated to explain around one half of the reductions in CHD mortality observed in these countries during these periods. In Finland, most of the decline in CHD

mortality observed between 1972 and 1992 could be explained by changes in three main coronary risk factors (serum total cholesterol, blood pressure and cigarette smoking),[544] while in the Seven Countries Study, CHD death rates during the latter part of the study were largely explained by changes in blood cholesterol levels during the early phases of the study.[587] Most recently, an analysis of CHD mortality rates in England and Wales between 1981 and 2000 estimated that 67% of the observed reduction in CHD mortality over that period could be explained by secular trends in cigarette smoking, blood cholesterol and blood pressure (though adverse trends in physical activity, obesity and diabetes resulted in the combined contribution of all risk factors being only 58%).[588]

In all these studies, the observation that blood lipids, blood pressure and cigarette smoking explain the vast majority of the combined effects attributable to changes in coronary risk factors (including, in some cases,[548;580] effects attributed to unknown factors) provides further supporting evidence regarding the combined importance of these factors. Furthermore, though the reduction in cigarette smoking rates observed in these studies was fairly large (30–40%), with the exception of the Finnish study,[544] the observed trends in blood cholesterol and blood pressure in these studies were small (3–7% reduction in average total cholesterol and 2–9% reduction in average blood pressure).[548;580;586;588] Therefore, had larger population–wide reductions in blood cholesterol and blood pressure occurred, substantially more CHD events would have been prevented. However, it is the ability of even small downwards trends in cholesterol and blood pressure to make notable contributions to reduced CHD mortality that confirms their importance, in addition to the more clearly demonstrated benefits from reducing cigarette smoking rates.

10.4.5 Implications for social inequalities in CHD

Motivated by the observation that despite decreases in CHD across all social class groups, relative differences between those at the top and those at the bottom of the spectrum are increasing, recent CHD prevention policies in the United Kingdom have placed an emphasis on reducing social inequalities in CHD.[54;55] The findings from the BRHS (chapter 9) are particularly relevant to the potential effects of these policies since the social class distribution of the men at study entry was close to that of all British men aged 35–64 at that time.[589] Furthermore, the estimates of social class differences were very similar to those observed in national CHD mortality statistics over a similar period (1981-1992).[590]

The key findings regarding social class and CHD from the work in this thesis are that:

1. Social class differences in CHD are marginally higher that would be estimated from baseline assessments of social class – manual social classes have approximately 40% higher risks of major CHD than non–manual social classes.

2. Approximately two–fifths of this difference in risk can be explained by differences in the adult coronary risk factors, with a further one fifth explained by influences in early–life.

3. Approximately 19% of all major CHD can be attributed to the excess risks of manual social classes, so that if manual social classes had experienced the same risks as non–manual social classes, 19% of CHD events would have been prevented.

4. If manual and non–manual social classes had had the same levels of the established adult coronary risk factors, then only around 10% of all major CHD events could have been attributed to the (remaining) excess risks of manual social classes.

The implications of these findings are that even if the reasons for social inequalities in CHD were understood entirely (and the risks could subsequently be reversed), only one in five premature major CHD events would be prevented in the long–term. Though this effect on population levels of CHD is appreciable (it is twice as high as the potential effectiveness of the most aggressive "high–risk" approaches advocated in the UK), it is still fairly small in comparison with the estimated reductions in CHD following population–wide reductions in total cholesterol and blood pressure (where at least one in four events could be prevented if mean total cholesterol and blood pressure could be lowered by just 5%). The limited contribution of reducing social inequalities reflects the high CHD risk and the unfavourable coronary risk profile of non–manual men in this study. The findings further indicate that the potential impact of any unidentified factors predisposing manual workers to higher risks of CHD than non–manual workers is likely to be small compared with the impact of targeting known key risk factors across all social classes. Although these conclusions only apply directly to the period 1980–2000 on which the data are based, they are still likely to be relevant to coronary heart disease prevention in the early twenty–first century. Current patterns of blood cholesterol and blood pressure (as measured in the Health Survey for England)[565] show less evidence of a social class gradient than was

observed in the British Regional Heart Study at baseline in 1978–1980, so that the case for population–wide prevention by reducing total cholesterol and blood pressure levels remains compelling. However, specific targeting of smoking cessation and physical activity among lower social class groups would clearly enhance the effectiveness of the population approach to prevention.

10.4.6 Recommendations

The analyses presented in this thesis clearly show that the effectiveness of the "high–risk" approach to the primary prevention of CHD is likely to be extremely limited unless used very much more widely than is currently recommended (particularly as recommended in the UK). Over one third of the middle–aged male population without pre–existing CHD would need to be treated with all four drugs to obtain benefits comparable with those following population–wide reductions in total cholesterol of 0.6 mmol/l and population–wide reductions in systolic blood pressure of 15 mmHg. This scale of prescribing would be consistent with the recently published Third European Joint Task Force report on cardiovascular disease prevention,[417] which recommends that priority should be given to individuals whose ten–year risk of fatal CVD (estimated from the SCORE project)[418] is at least 5% – under this criterion, 36% of the BRHS men would be defined as being at "high-risk" at baseline. However, treating such a large proportion of the "healthy" population would have considerable financial implications with pharmacological high–risk approaches becoming less cost effective as the absolute risk threshold is lowered. In comparison, population approaches have been shown to be highly cost effective,[435] and, more importantly, focus on the determinants of risk factor distributions rather than simply the treatment of risk factors. Population approaches focused on blood lipids, blood pressure and cigarette smoking have been demonstrated to be feasible and effective; analysis of time–trend studies have demonstrated the important effects that even small population–wide changes in these factors can have. Population approaches may be more likely to reduce the development of atherosclerosis, while a high–risk strategy implemented without priority given to population approaches would ensure a steady supply of middle–aged people requiring drug treatment.

The results emphasize the considerable potential benefits of population–wide strategies for CHD prevention. In the United Kingdom, mean total cholesterol and blood pressure

levels remain high by international standards, and though the Health Survey for England does suggest that important reductions in mean cholesterol levels have been achieved in recent years,[545] this is difficult to validate because of a change the laboratory methods used during this time. Current public health policy for CHD prevention in the United Kingdom gives little emphasis to the importance of reducing total cholesterol and blood pressure levels in the population, nor to the crucial role of Governmental action likely to be necessary to bring about such changes (for instance legislation to decrease salt and fat content in processed foods). Spending on such policies currently comprises only 1–2% of total spending on health in many Western countries including the United Kingdom.[591;592] It is likely that by giving greater priority to population–wide reductions in blood cholesterol and blood pressure, and continued emphasis to reducing cigarette smoking, the substantial gains in CHD prevention already achieved over the last two decades may be maintained, particularly in the face of adverse gradients in sedentary behaviour, obesity and diabetes.

Appendix A

Publications from this thesis

1. Emberson JR, Whincup PH, Morris RW, Walker M, Lowe GDO, Rumley A. Extent of regression dilution for established and novel coronary risk factors: results from the British Regional Heart Study. *Eur J Cardiovasc Prev Rehabil* 2004:**11**; 125–134

2. Emberson JR, Whincup PH, Morris RW, Walker M. Re–assessing the contribution of serum total cholesterol, blood pressure and cigarette smoking to the aetiology of coronary heart disease: impact of regression dilution bias. *Eur Heart J* 2003:**24**; 1719–1726

3. Emberson JR, Whincup PH, Morris RW, Walker M, Ebrahim S. Evaluating the impact of population and high–risk strategies for the primary prevention of cardiovascular disease. *Eur Heart J* 2004:**25**; 484–491

4. Emberson JR, Whincup PH, Morris RW, Walker M. Social class differences in coronary heart disease in middle–aged British men: implications for prevention. *Int J Epidemiol* 2004:**33**; 289–296

5. Emberson JR, Whincup PH, Morris RW, Walker M. Author's response: Reducing social inequalities and the prevention of coronary heart disease (letter). *Int J Epidemiol* 2004:**33**; 1152–1153

6. Emberson JR, Whincup PH, Morris RW, Wannamethee SG, Shaper AG. Lifestyle and cardiovascular disease in middle–aged British men: the effect of adjusting for within–person variation. *Eur Heart J* 2005**(in press)**

7. Emberson JR, Shaper AG, Wannamethee SG, Morris PH, Whincup PH. Alcohol intake in middle–age and risk of cardiovascular disease and mortality: accounting for intake variation over time. *Am J Epidemiol* 2005**(in press)**

Extent of regression dilution for established and novel coronary risk factors: results from the British Regional Heart Study

Jonathan R. Emberson[a], Peter H. Whincup[b], Richard W. Morris[a], Mary Walker[a], Gordon D.O. Lowe[c] and Ann Rumley[c]

1741-8267 DOI: 10.1097/01.hjr.0000114967.39211.e5

European Heart Journal (2003) 24, 1719–1726

Clinical research

Re-assessing the contribution of serum total cholesterol, blood pressure and cigarette smoking to the aetiology of coronary heart disease: impact of regression dilution bias

Jonathan R. Emberson[a*], Peter H. Whincup[b], Richard W. Morris[a], Mary Walker[a]

0195-668X/03/$ - see front matter
doi:10.1016/S0195-668X(03)00471-8

European Heart Journal (2004) 25, 484–491

Clinical research

Evaluating the impact of population and high-risk strategies for the primary prevention of cardiovascular disease

Jonathan Emberson[a,*], Peter Whincup[b], Richard Morris[a], Mary Walker[a], Shah Ebrahim[c]

0195-668X/$ - see front matter
doi:10.1016/j.ehj.2003.11.012

International Journal of Epidemiology 2004;**33**:289–296
DOI: 10.1093/ije/dyh006

Social class differences in coronary heart disease in middle-aged British men: implications for prevention

Jonathan R Emberson,[1] Peter H Whincup,[2] Richard W Morris[1] and Mary Walker[1]

Appendix B

Baseline questionnaire

The baseline questionnaire completed by study participants is reproduced on the following pages. For information regarding the analysis and coding of this questionnaire refer to chapter 4, section 4.3.1.

1

Serial Number [ ][ ][ ][ ][ ][ ][ ][ ] 1

Card Number [0][1] 9

Date of Screening [ ][ ][ ][ ][ ][ ] 11

Time of Screening [ ][ ][ ][ ] 17

1. GENERAL

What is your date of birth?

Day [ ][ ] 21

Month [ ][ ] 23

Year 19 [ ][ ] 25

Where were you born?

Town .....

County .....

Country .....

1.2 How many years have you lived within 10 miles of this town?
If you have moved to this area within the last five years, where did you move from?
.....
[ ][ ] 27 years

1.3 What is your marital status?

Single 1
Married 2
Widowed 3
Other 4

[ ] 29

1.4 How many children do you have?

	M	F	
<5 yrs	[ ]	[ ]	30
5-10 yrs.	[ ]	[ ]	32
11-16 yrs.	[ ]	[ ]	34
> 16 yrs.	[ ]	[ ]	36

2 YOUR FATHER

2.1 Where was your Father born?

Town .....

County .....

Country .....

2.2 Is your father alive? (Y/N) [ ] 38

2.3 How old is he now? / How old was he when he died? [ ][ ] 39 years

2.4 If your father has died, what were you told was the cause of his death?

Heart trouble 1
High blood pressure 2
Stroke 3
Respiratory disease 4
Cancer of lung 5
Other cancer 6
Accident or injury 7
Other 8
Don't know 9

[ ] 41

3 YOUR MOTHER

3.1 Where was your mother born?

Town .....

County .....

Country .....

3.2 Is your mother alive? (Y/N) [ ] 42

3.3 How old is she now? / How old was she when she died? [ ][ ] years 43

3.4 If your mother has died, what were you told was the cause of her death?

Heart trouble 1
High blood pressure 2
Stroke 3
Respiratory disease 4
Cancer of breast 5
Other cancer 6
Accident or injury 7
Other 8
Don't know 9

[ ] 45

4. OCCUPATION

4.1 What is your present job? .....

If employed go to question 4.4

4.2 If you are unemployed, for how long has this been?

<6weeks 1
6wk.-5mo. 2
6mo. –1yr. 3
> 1 year 4

[ ] 46

4.3 Is this because of ill health? (Y/N) ☐ 47

.....

4.4 What kind of work have you done for the longest period of time?

.....

4.5 What business or industry is this?

.....

4.6 How many years have you done this kind of work? ☐☐ years 48

4.7 Are / were you:

SELF-EMPLOYED	with 25 or more employees	1	☐ 50
	with less than 25 employees	2	
	without employees	3	
MANAGER	of 25 or more people	4	
	of less than 25 people	5	
FOREMAN	.....	6	
ORDINARY EMPLOYEE	.....	7	
ARMED SERVICES	.....	8	

5 SEVERE CHEST PAIN

5.1 Have you ever had a severe pain in your chest lasting for half an hour or more? (Y/N) ☐ 51

If NO, go to question 6.

5.2 Where did you get this severe pain? ☐ 52

(Show chart.)

5.3 Did you see a doctor because of this pain? (Y/N) ☐ 55

6 CHEST PAIN

6.1 Do you ever have any pain or discomfort in your chest? (Y/N) ☐ 56

If NO, go to question 7.

6.2 When last did you get the pain? ☐ 57

Within 1 month	1
1-5 months ago	2
6-12 months ago	3
Over 1 year ago	4
Occasionally	5

6.3 How often do you get it? ☐ 58

Daily	1
Weekly	2
Monthly	3
Once only	4
Occasionally	5

6.4 Where do you get this pain or discomfort? ☐ 59

(Show chart.)

6.5 When you walk at an ordinary pace on the level, does this produce the pain? (Y/N) ☐ 62

6.6 When you walk uphill or hurry, does this produce the pain? (Y/N) ☐ 63

6.7 When you get any pain or discomfort in your chest on walking, what do you do? ☐ 64

Stop	1
Slow down	2
Continue at the same pace	3

6.8 Does the pain or discomfort in your chest go away if you stand still? (Y/N) ☐ 65

6.9 How long does it take to go away?

10 minutes or less	1
more than 10 minutes	2

☐ 66

7.0 PHLEGM, COUGH AND BREATHING

7.1 Do you usually bring up phlegm (spit) from your chest first thing in the morning in the winter? (Y/N) ☐ 67

If NO, go to question 7.4

7.2 Do you bring up phlegm like this on most days for as much as 3 months in the winter each year? (Y/N) ☐ 68

7.3 In the past 3 years have you ever had a period of increased cough and phlegm lasting 3 weeks or more? ☐ 69

Yes, once	1
Yes, twice or more	2
Never	3

7.4 Does your chest sound wheezy or whistling on most days (or nights)? (Y/N) ☐ 70

7.5 Does the weather affect your breathing?
And if so, at what season of the year is it most affected?

Not affected	1	71
Winter	2	
Summer	3	
Both	4	

8 BREATHLESSNESS

8.1 Do you get short of breath walking with people your own age on level ground? (Y/N) 72

8.2 On walking up hills or stairs, do you get more breathless than people you on age? (Y/N) 73

8.3 Do you ever have to stop walking because of breathlessness? (Y/N) 74

S.E.G. 75

Social class 77

Activity score 78

6

Serial Number

Card Number 0 2

9 LEG PAIN

9.1 Do you ever get pain in your calf muscles on walking at an ordinary pace, on the level? (Y/N) 11

9.2 Do you get pain in your calf muscles when you walk uphill or hurry? (Y/N) 12

10 MEDICAL HISTORY

10.1 Have you ever been told by a doctor that you have, or have had, any of the following?

Angina	(Y/N)	13
Heart attack	(Y/N)	14
Coronary thrombosis	(Y/N)	15
Myocardial infarction	(Y/N)	16
Other heart trouble	(Y/N)	17
High blood pressure	(Y/N)	18
Stroke	(Y/N)	19
Diabetes	(Y/N)	20
Peptic ulcer	(Y/N)	21
Gout	(Y/N)	22
Gall bladder disease	(Y/N)	23
Thyroid disease	(Y/N)	24
Arthritis	(Y/N)	25
Bronchitis	(Y/N)	26
Asthma	(Y/N)	27
Other condition (s) including surgery.....	(Y/N)	28

10.2 Are you on any regular medical treatment from a doctor for any condition? (Y/N) 29

If NO, go to question 10.3

Do you know if the pills / medicines /injections are:-

Tranquillizers	Y/N	30
Pain killers	Y/N	31
Antihypertensive drugs	Y/N	32
Anti coagulants	Y/N	33
Lipid lowering drugs	Y/N	34

Oral antidiabetics	Y/N		35
Injection of insulin	Y/N		36
Any others	Y/N		37
Don't know	Y/N		38

10.3 Have you taken any of these in the last 48 hours?

Tranquillizers	Y/N		39
Pain killers	Y/N		40
Antihypertensive drugs	Y/N		41
Anti coagulants	Y/N		42
Lipid lowering drugs	Y/N		43
Oral antidiabetics	Y/N		44
Injection of insulin	Y/N		45
Any others	Y/N		46
Don't know	Y/N		47

11 DIET & ALCOHOL

11.1 How many times during an average week would you have the following foods?

Meat (including beef, lamb, pork, bacon in any form)		48
Chicken		50
Fish		52
Eggs - how many eggs do you eat in a week		54
Cheese – how often do you eat cheese, including cheese dishes?		56
Breakfast cereals – how often do you eat these (porridge included)? State kind		58

11.2 What kinds of bread do you eat ?

White	Y/N		60
Brown	Y/N		61
Wholemeal	Y/N		62
Other	Y/N		63

11.3 Spreading fats: What kinds do you use at home?

Butter	Y/N		64
Margarine	Y/N		65

(State kind or brand name.)

11.4 Do you take sugar?

In tea	Y/N		86
In coffee	Y/N		87
In other drinks	Y/N		88

11.5 Do you use milk?

On cereals	Y/N		69
In tea	Y/N		70
In coffee	Y/N		71
As a milk drink	Y/N		72

11.6 (i) Would you describe your present alcohol intake as:

None	1		
On special occasions only	2		
Once or twice a month	3		73
Weekends	4		
Daily / most days	5		

If NONE, go to question 12

(ii) What type of drink do you usually take?

Beer	1		
Spirits	2		
Wine/sherry	3		74
Mixed beer & spirits	4		
Mixed beer, spirits, wine and sherry	5		

(iii) How much do you usually take?

2 drinks a day or less	1		
3-6 drinks a day	2		
more than 6 drinks a day	3		75

(One drink is a single whisky, gin or brandy, a glass of wine, sherry or port or half a pint of beer.)

Serial Number

Card Number 0 3

12 SMOKING

12.1 (i) Do you smoke at present?

Yes, regularly 1
No 2
Occasionally 3

11

If NO, go to question 12.6

(ii) How old were you when you started? years 12

(iii) Have you ever given up smoking? (Y/N) 14

(iv) If yes, what is the maximum time for which you have given up smoking? years 15

12.2 (i) Do you smoke cigarettes now?

Yes regularly 1
No 2
Occasionally (<1 day) 3

17

If NO, or OCCASIONALLY go to question 12.3

(ii) How many cigarettes do you usually smoke a day? 18

(iii) If hand rolled, how much tabacco do you use a week? (ozs.) ozs. 20

Now proceed to 12.4

12.3 (i) Were you previously a regular cigarette smoker? (Y/N) 22

(ii) If Yes, how many cigarettes did you usually smoke a day? 23

(iii) At what age did you change to a pipe and / or cigars? years 25

12.4 (i) Do you smoke a pipe now?

Yes regularly 1
No 2
Occasionally 3

27

If NO or OCCASIONALLY go to question 12

(ii) If YES, how many ozs. a week do you smoke? ozs. 20

12.5 (i) Do you smoke a pipe now?

Yes regularly 1
No 2
Occasionally 3

30

(ii) If YES, how many cigars do you smoke a day?

Large 31
Small 32

If you smoke ANYTHING currently, go to question 13.

12.6 (i) Have you ever smoked for a more than 1 month ? (Y/N) 35

How much did you usually smoke

Cigarettes (per day) 36 38
Pipe (ozs) (per week) 40
Cigars (per day) Large 42
Small

If NO, go to question 13.

(ii) At what age did you start smoking? years 44

(iii) At what age did you finally stop smoking? years 46

(iv) What was the maximum time between these two ages for which you gave up smoking? years 48

13 EXERCISE

13.1 (i) Do you usually walk or cycle in the course of your journeys to or from work each day?

No 1
Walk 2
Cycle 3

50

If YES, how many minutes do these journeys take? mins 51

(ii) Apart from your journeys to or from work, do you usually walk or cycle on weekdays?

No 1
Walk 2
Cycle 3

50

If YES, how many minutes do you walk/cycle each day? mins 51

(iii) Would you say that in your occupation you are physically :

Very active 1
Fairly active 2
Average 3
Fairly inactive 4
Very inactive 5

58

13.2 On average, a man of your age spends 4 hours on most weekends on some of the following activities: walking, gardening, household chores, DIY projects. Compared to such a man, how physically active do you consider yourself?

Very active 1
Fairly active 2
Average 3
Fairly inactive 4
Very inactive 5

57

13.3 Apart from these activities, do you take active physical exercise, e.g. running, digging, swimming, tennis, golf, sailing, etc.

No	1	
Occassionally	2	☐ 58
Frequently	3	

If NO or Occasionally – stop here.

13.4 Please state type of activity.....

13.5 How many years have you been involved in this activity?

☐☐ years 59

13.6 How many times a month (on average) do you undertake these activities?

Winter	☐☐	61
Summer	☐☐	63

Administrator	☐	65
Coder	☐	66

References

[1] The World Health Report 2002. Reducing risks, promoting healthy lifestyle. Geneva: World Health Organisation, 2002.

[2] Reddy KS and Yusuf S. Emerging epidemic of cardiovascular disease in developing countries. *Circulation*, 1998; **97**:596–601.

[3] British Heart Foundation Statistics Database. Coronary Heart Disease Statistics. British Heart Foundation, 2002.

[4] Wood D, De Backer G, Faergeman O, Graham I, Mancia G, and Pyorala K. Prevention of coronary heart disease in clinical practice: recommendations of the Second Joint Task Force of European and other Societies on Coronary Prevention. *Atherosclerosis*, 1998; **140**:199–270.

[5] Primatesta P. Health Survey for England. Volume 1: findings, Chapter 2 – Prevalence of Cardiovascular Disease. London: The Stationery Office, 1998.

[6] Lampe FC, Morris RW, Whincup PH, Walker M, Ebrahim S, and Shaper AG. Is the prevalence of coronary heart disease falling in British men? *Heart*, 2001; **86**:499–505.

[7] Yusuf S, Reddy S, Ounpuu S, and Anand S. Global burden of cardiovascular diseases: part I: general considerations, the epidemiologic transition, risk factors, and impact of urbanization. *Circulation*, 2001; **104**:2746–2753.

[8] Yusuf S, Reddy S, Ounpuu S, and Anand S. Global burden of cardiovascular diseases: Part II: variations in cardiovascular disease by specific ethnic groups and geographic regions and prevention strategies. *Circulation*, 2001; **104**:2855–2864.

[9] Reddy KS. Cardiovascular diseases in India. *World Health Statistics Quarterly*, 1993; **46**:101–107.

[10] Weissberg PL. Coronary disease: Atherogenesis: current understanding of the causes of atheroma. *Heart*, 2000; **83**:247–252.

[11] Davies MJ and Thomas AC. Plaque fissuring–the cause of acute myocardial infarction, sudden ischaemic death, and crescendo angina. *Br Heart J*, 1985; **53**:363–373.

[12] Kannel WB and Larson M. Long–term epidemiologic prediction of coronary disease. The Framingham experience. *Cardiology*, 1993; **82**:137–152.

[13] Devereux RB and Alderman MH. Role of preclinical cardiovascular disease in the evolution from risk factor exposure to development of morbid events. *Circulation*, 1993; **88**:1444–1455.

[14] Henderson A. Coronary heart disease: overview. *Lancet*, 1996; **348**:s1–s4.

[15] Fuller WA and Hidiroglou MA. Regression estimates after correcting for attenuation. *J Am Stat Assoc*, 1978; **73**:99–104.

[16] Rosner B, Willett WC, and Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within–person measurement error. *Stats Med*, 1989; **8**:1051–1069.

[17] MacMahon S, Peto R, Cutler J, Collins R, Sorlie P, Neaton J, Abbott R, Godwin J, Dyer A, and Stamler et al. Blood pressure, stroke, and coronary heart disease. Part 1, Prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet*, 1990; **335**:765–774.

[18] de Klerk NH, English DR, and Armstrong BK. A review of the effects of random measurement error on relative risk estimates in epidemiological studies. *Int J Epidemiol*, 1989; **18**:705–712.

[19] Clarke R, Shipley M, Lewington S, Youngman L, Collins R, Marmot M, and Peto R. Underestimation of risk associations due to regression dilution in long–term follow–up of prospective studies. *Am J Epidemiol*, 1999; **150**:341–353.

[20] Frost C and Thompson SG. Correcting for regression dilution bias: comparison of methods for a single predictor variable. *JRSS (A)*, 2000; **163**:173–189.

[21] Bashir SA and Duffy SW. The correction of risk estimates for measurement error. *Ann Epidemiol*, 1997; **7**:154–164.

[22] Bashir SA, Duffy SW, and Qizilbash N. Repeat measurement of case–control data: corrections for measurement error in a study of ischaemic stroke and haemostatic factors. *Int J Epidemiol*, 1997; **26**:64–70.

[23] Duffy SW, Rohan TE, and Day NE. Misclassification in more than one factor in a case–control study: a combination of Mantel–Haenszel and maximum likelihood approaches. *Stats Med*, 1989; **8**:1529–1536.

[24] Qizilbash N, Duffy SW, and Rohan TE. Repeat measurement of case–control data: correcting risk estimates for misclassification due to regression dilution of lipids in transient ischemic attacks and minor ischemic strokes. *Am J Epidemiol*, 1991; **133**:832–838.

[25] Knuiman MW, Divitini ML, Buzas JS, and Fitzgerald PE. Adjustment for regression dilution in epidemiological regression analyses. *Ann Epidemiol*, 1998; **8**:56–63.

[26] Rosner B, Spiegelman D, and Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. *Am J Epidemiol*, 1990; **132**:734–745.

[27] Phillips AN and Smith GD. How independent are 'independent' effects? Relative risk estimation when correlated exposures are measured imprecisely. *J Clinical Epidemiol*, 1991; **44**:1223–1231.

[28] Kannel WB. The Framingham Study: Its 50–year legacy and future promise. *J Atheroscler Throm*, 2000; **6**:60–66.

[29] Keys AB. Seven Countries: a multivariate analysis of death and coronary heart disease. Cambridge, Mass: Harvard University Press, 1980.

[30] Corday E and Corday SR. Editorial: Prevention of heart disease by control of risk factors: the time has come to face the facts. *Am J Cardiol*, 1975; **35**:330–333.

[31] Wilmshurst P. Temperature and cardiovascular mortality. *BMJ*, 1994; **309**:1029–1030.

[32] Frank JW. Why 'population health'? *Can J Pub Health*, 1995; **86**:162–164.

[33] Syme SL. Rethinking disease: where do we go from here? *Ann Epidemiol*, 1996; **6**:463–468.

[34] Kaplan GA and Keil JE. Socioeconomic factors and cardiovascular disease: a review of the literature. *Circulation*, 1993; **88**:1973–1998.

[35] Nieto FJ. Cardiovascular disease and risk factor epidemiology: a look back at the epidemic of the 20th century. *Am J Public Health*, 1999; **89**:292–294.

[36] Braunwald E. Shattuck lecture–cardiovascular medicine at the turn of the millennium: triumphs, concerns, and opportunities. *N Engl J Med*, 1997; **337**:1360–1369.

[37] Hennekens CH. Increasing burden of cardiovascular disease: current knowledge and future directions for research on risk factors. *Circulation*, 1998; **97**:1095–1102.

[38] Lefkowitz RJ and Willerson JT. Prospects for cardiovascular research. *JAMA*, 2001; **285**:581–587.

[39] Schnohr P, Jensen JS, Scharling H, and Nordestgaard BG. Coronary heart disease risk factors ranked by importance for the individual and community. A 21 year follow–up of 12 000 men and women from The Copenhagen City Heart Study. *Eur Heart J*, 2002; **23**:620–626.

[40] Jenkins CD. Epidemiology of cardiovascular diseases. *J Consult Clin Psychol*, 1988; **56**:324–332.

[41] Stamler J, Dyer AR, Shekelle RB, Neaton J, and Stamler R. Relationship of baseline major risk factors to coronary and all–cause mortality, and to longevity: findings from long–term follow–up of Chicago cohorts. *Cardiology*, 1993; **82**:191–222.

[42] Magnus P and Beaglehole R. The real contribution of the major risk factors to the coronary epidemics: time to end the 'only–50%' myth. *Arch Int Med*, 2001; **161**:2657–2660.

[43] Beaglehole R and Magnus P. The search for new risk factors for coronary heart disease: occupational therapy for epidemiologists? *Int J Epidemiol*, 2002; **31**:1117–1122.

[44] Stamler J, Stamler R, Neaton JD, Wentworth D, Daviglus ML, Garside D, Dyer AR, Liu K, and Greenland P. Low risk–factor profile and long–term cardiovascular and noncardiovascular mortality and life expectancy: findings for 5 large cohorts of young adult and middle–aged men and women. *JAMA*, 1999; **282**:2012–2018.

[45] Stampfer MJ, Hu FB, Manson JE, Rimm EB, and Willett WC. Primary prevention of coronary heart disease in women through diet and lifestyle. *N Engl J Med*, 2000; **343**:16–22.

[46] Greenland P, Knoll MD, Stamler J, Neaton JD, Dyer AR, Garside DB, and Wilson PW. Major Risk Factors as Antecedents of Fatal and Nonfatal Coronary Heart Disease Events. *JAMA*, 2003; **290**:891–897.

[47] Khot UN, Khot MB, Bajzer CT, Sapp SK, Ohman EM, Brener SJ, Ellis SG, Lincoff AM, and Topol EJ. Prevalence of Conventional Risk Factors in Patients With Coronary Heart Disease. *JAMA*, 2003; **290**:898–904.

[48] Rose G. The strategy of preventive medicine. Oxford: Oxford University Press, 1992.

[49] Strachan D and Rose G. Strategies of prevention revisited: effects of imprecise measurement of risk factors on the evaluation of 'high–risk' and 'population–based' approaches to prevention of cardiovascular disease. *J Clinical Epidemiol*, 1991; **44**:1187–1196.

[50] Anderson KM, Wilson PW, Odell PM, and Kannel WB. An updated coronary risk profile. A statement for health professionals. *Circulation*, 1991; **83**:356–362.

[51] Yusuf S. Two decades of progress in preventing vascular disease. *Lancet*, 2002; **360**:2–3.

[52] Rose G. Sick individuals and sick populations. *Int J Epidemiol*, 1985; **14**:32–38.

[53] Marmot MG, Adelstein AM, Robinson N, and Rose GA. Changing social–class distribution of heart disease. *BMJ*, 1978; **2**:1109–1112.

[54] Sir Donald Acheson. Independent inquiry into inequalities in health. London: The Stationery Office, 2002.

[55] Secretary of State for Health. Saving Lives: Our Healthier Nation. London: The Stationery Office, 1999.

[56] Robertson WB. The international atherosclerosis project. *Pathologia Et Microbiologia*, 1967; **30**:810–816.

[57] Relationship of atherosclerosis in young men to serum lipoprotein cholesterol concentrations and smoking. A preliminary report from the Pathobiological Determinants of Atherosclerosis in Youth (PDAY) Research Group. *JAMA*, 1990; **264**:3018–3024.

[58] Rimm EB, Giovannucci EL, Willett WC, Colditz GA, Ascherio A, Rosner B, and Stampfer MJ. Prospective study of alcohol consumption and risk of coronary disease in men. *Lancet*, 1991; **338**:464–468.

[59] Colditz GA. The Nurses' Health Study: a cohort of US women followed since 1976. *J Am Med Wom Assoc*, 1972; **50**:40–44.

[60] The Multiple Risk Factor Intervention Trial (MRFIT). A national study of primary prevention of coronary heart disease. *JAMA*, 1976; **235**:825–827.

[61] Shaper AG, Pocock SJ, Walker M, Cohen NM, Wale CJ, and Thomson AG. British Regional Heart Study: cardiovascular risk factors in middle–aged men in 24 towns. *BMJ*, 1981; **283**:179–186.

[62] Reid DD, Brett GZ, Hamilton PJ, Jarrett RJ, Keen H, and Rose G. Cardiorespiratory disease and diabetes among middle–aged male Civil Servants. A study of screening and intervention. *Lancet*, 1974; **1**:469–473.

[63] Hawthorne VM, Watt GC, Hart CL, Hole DJ, Smith GD, and Gillis CR. Cardiorespiratory disease in men and women in urban Scotland: baseline characteristics of the Renfrew/Paisley (midspan) study population. *Scott Med J*, 1995; **40**:102–107.

[64] Tunstall-Pedoe H, Smith WC, Crombie IK, and Tavendale R. Coronary risk factor and lifestyle variation across Scotland: results from the Scottish Heart Health Study. *Scott Med J*, 1989; **34**:556–560.

[65] Ducimetie're P, Richard JL, Cambien F, Rakotovao R, and Claude JR. Coronary heart disease in middle–aged Frenchmen. Comparisons between Paris Prospective Study, Seven Countries Study, and Pooling Project. *Lancet*, 1980; **1**:1346–1350.

[66] Leren P, Askevold EM, Foss OP, Froili A, Grymyr D, Helgeland A, Hjermann I, Holme I, Lund-Larsen PG, and Norum et al. The Oslo study. Cardiovascular disease in middle–aged and young Oslo men. *Acta Medica Scandinavica*, 1975; **588**:1–38.

[67] Yuan JM, Ross RK, Gao YT, and Yu MC. Body weight and mortality: a prospective evaluation in a cohort of middle–aged men in Shanghai, China. *Int J Epidemiol*, 1998; **27**:824–832.

[68] Garcia-Palmieri MR, Feliberti M, Costas R Jr, Colon AA, Cruz-Vidal M, Cortes-Alicea M, Ayala AM, Sobrino R, and Torres et al. An epidemiological study on coronary heart disease in Puerto Rico: The Puerto Rico Heart Health Program. *Boletin De La Asociacion Medica De Puerto Rico*, 1969; **61**:174–179.

[69] Prevention of Coronary Heart Disease. Report of a WHO Expert Committee. WHO technical report series 678. Geneva: World Health Organization, 1982.

[70] Law MR, Wald NJ, and Thompson SG. By how much and how quickly does reduction in serum cholesterol concentration lower risk of ischaemic heart disease? *BMJ*, 1994; **308**:367–372.

[71] Law MR and Wald NJ. Risk factor thresholds: their existence under scrutiny. *BMJ*, 2002; **324**:1570–1576.

[72] Stamler J, Wentworth D, and Neaton JD. Is relationship between serum cholesterol and risk of premature death from coronary heart disease continuous and graded? Findings in 356,222 primary screenees of the Multiple Risk Factor Intervention Trial (MRFIT). *JAMA*, 1986; **256**:2823–2828.

[73] Law MR, Wald NJ, and Rudnicka AR. Quantifying effect of statins on low density lipoprotein cholesterol, ischaemic heart disease, and stroke: systematic review and meta–analysis. *BMJ*, 2003; **326**:1423–1427.

[74] WHO cooperative trial on primary prevention of ischaemic heart disease with clofibrate to lower serum cholesterol: final mortality follow–up. Report of the Committee of Principal Investigators. *Lancet*, 1984; **2**:600–604.

[75] Strandberg TE, Salomaa VV, Naukkarinen VA, Vanhanen HT, Sarna SJ, and Miettinen TA. Long–term mortality after 5–year multifactorial primary prevention of cardiovascular diseases in middle–aged men. *JAMA*, 1991; **266**:1225–1229.

[76] Muldoon MF, Manuck SB, and Matthews KA. Lowering cholesterol concentrations and mortality: a quantitative review of primary prevention trials. *BMJ*, 1990; **301**:309–314.

[77] Davey Smith G and Pekkanen J. Should there be a moratorium on the use of cholesterol lowering drugs? *BMJ*, 1992; **304**:431–434.

[78] Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high–risk individuals: a randomised placebo–controlled trial. *Lancet*, 2002; **360**:7–22.

[79] Sacks FM, Pfeffer MA, Moye LA, Rouleau JL, Rutherford JD, Cole TG, Brown L, Warnica JW, Arnold JM, and Wun et al. The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. Cholesterol and Recurrent Events Trial investigators. *N Engl J Med*, 1996; **335**:1001–1009.

[80] Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet*, 1994; **344**:1383–1389.

[81] The Long-Term Intervention with Pravastatin in Ischaemic Disease (LIPID) Study Group. Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. *N Engl J Med*, 1998; **339**:1349–1357.

[82] Shepherd J, Cobbe SM, Ford I, Isles CG, Lorimer AR, MacFarlane PW, McKillop JH, and Packard CJ. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. *N Engl J Med*, 1995; **333**:1301–1307.

[83] Downs JR, Clearfield M, Weis S, Whitney E, Shapiro DR, Beere PA, Langendorfer A, Stein EA, Kruyer W, and Gotto et al. Primary prevention of acute coronary events with lovastatin in men and women with average cholesterol levels: results of AFCAPS/TexCAPS. Air Force/Texas Coronary Atherosclerosis Prevention Study. *JAMA*, 1998; **279**:1615–1622.

[84] LaRosa JC, He J, and Vupputuri S. Effect of statins on risk of coronary disease: a meta–analysis of randomized controlled trials. *JAMA*, 1999; **282**:2340–2346.

[85] Ballantyne CM. Low–density lipoproteins and risk for coronary artery disease. *Am J Cardiol*, 1998; **82**:3Q–12Q.

[86] Rosenson RS and Tangney CC. Antiatherothrombotic properties of statins: implications for cardiovascular event reduction. *JAMA*, 1998; **279**:1643–1650.

[87] Vaughan CJ, Murphy MB, and Buckley BM. Statins do more than just lower cholesterol. *Lancet*, 1996; **348**:1079–1082.

[88] Sacks FM, Moye LA, Davis BR, Cole TG, Rouleau JL, Nash DT, Pfeffer MA, and Braunwald E. Relationship between plasma LDL concentrations during treatment with pravastatin and recurrent coronary events in the Cholesterol and Recurrent Events trial. *Circulation*, 1998; **97**:1446–1452.

[89] Simes RJ, Marschner IC, Hunt D, Colquhoun D, Sullivan D, Stewart RAH, Hague W, Keech A, Thompson P, and White H. Relationship between lipid levels and clinical outcomes in the Long–term Intervention with Pravastatin in Ischemic Disease (LIPID) Trial: to what extent is the reduction in coronary events with pravastatin explained by on–study lipid levels? *Circulation*, 2002; **105**:1162–1169.

[90] Kagan A, McGee DL, Yano K, Rhoads GG, and Nomura A. Serum cholesterol and mortality in a Japanese–American population: the Honolulu Heart program. *Am J Epidemiol*, 1981; **114**:11–20.

[91] Sorlie PD and Fienleib M. The serum cholesterol–cancer relationship: an analysis of time trends in the Framingham Study. *J Natl Cancer Inst*, 1982; **69**:989–996.

[92] Cambien F, Ducimetiere P, and Richard J. Total serum cholesterol and cancer mortality in a middle–aged male population. *Am J Epidemiol*, 1980; **112**:388–394.

[93] International Collaborative Group. Circulating cholesterol level and risk of death from cancer in men aged 40 to 69 years. *JAMA*, 1982; **248**:2853–2859.

[94] Sherwin RW, Wentworth DN, Cutler JA, Hulley SB, Kuller LH, and Stamler J. Serum cholesterol levels and cancer mortality in 361,662 men screened for the Multiple Risk Factor Intervention Trial. *JAMA*, 1987; **257**:943–948.

[95] Knekt P, Reunanen A, Aromaa A, Heliovaara M, Hakulinen T, and Hakama M. Serum cholesterol and risk of cancer in a cohort of 39,000 men and women. *J Clinical Epidemiol*, 1988; **41**:519–530.

[96] Wannamethee G, Shaper AG, Whincup PH, and Walker M. Low serum total cholesterol concentrations and mortality in middle aged British men. *BMJ*, 1995; **311**:409–413.

[97] Davey Smith G, Shipley MJ, Marmot MG, and Rose G. Plasma cholesterol concentration and mortality. The Whitehall Study. *JAMA*, 1992; **267**:70–76.

[98] Isles CG, Hole DJ, Gillis CR, Hawthorne VM, and Lever AF. Plasma cholesterol, coronary heart disease, and cancer in the Renfrew and Paisley survey. *BMJ*, 1989; **298**:920–924.

[99] Morris DL, Borhani NO, Fitzsimons E, Hardy RJ, Hawkins CM, Kraus JF, Labarthe DR, Mastbaum L, and Payne GH. Serum cholesterol and cancer in the Hypertension Detection and Follow–up Program. *Cancer*, 1983; **52**:1754–1759.

[100] Schatzkin A, Hoover RN, Taylor PR, Ziegler RG, Carter CL, Larson DB, and Licitra LM. Serum cholesterol and cancer in the NHANES I epidemiologic followup study. National Health and Nutrition Examination Survey. *Lancet*, 1987; **2**:298–301.

[101] Garcia-Palmieri MR, Sorlie PD, Costas R Jr, and Havlik RJ. An apparent inverse relationship between serum cholesterol and cancer mortality in Puerto Rico. *Am J Epidemiol*, 1981; **114**:29–40.

[102] Law MR, Thompson SG, and Wald NJ. Assessing possible hazards of reducing serum cholesterol. *BMJ*, 1994; **308**:373–379.

[103] Simes J. Long–term effectiveness and safety of pravastatin in 9014 patients with coronary heart disease and average cholesterol concentrations: the LIPID trial follow–up. *Lancet*, 2002; **359**:1379–1387.

[104] Clarke R, Frost C, Collins R, Appleby P, and Peto R. Dietary lipids and blood cholesterol: quantitative meta–analysis of metabolic ward studies. *BMJ*, 1997; **314**:112–117.

[105] Hegsted DM, Ausman LM, Johnson JA, and Dallal GE. Dietary fat and serum lipids: an evaluation of the experimental data. *Am J Clin Nutr*, 1993; **57**:875–883.

[106] Hegsted DM, McGandy RB, Myers ML, and Stare FJ. Quantitative effects of dietary fat on serum cholesterol in man. *Am J Clin Nutr*, 1965; **17**:281–295.

[107] Keys A, Anderson JT, and Grande F. Prediction of serum–cholesterol responses of man to changes in fats in the diet. *Lancet*, 1957; **270**:959–966.

[108] Dattilo AM and Kris-Etherton PM. Effects of weight reduction on blood lipids and lipoproteins: a meta–analysis. *Am J Clin Nutr*, 1992; **56**:320–328.

[109] Ripsin CM, Keenan JM, Jacobs DR Jr, Elmer PJ, Welch RR, Van Horn L, Liu K, Turnbull WH, and Thye et al. Oat products and lipid lowering. A meta–analysis. *JAMA*, 1992; **267**:3317–3325.

[110] Silagy C and Neil A. Garlic as a lipid lowering agent–a meta–analysis. *J R Coll Phys Lond*, 1994; **28**:39–45.

[111] Harris WS. Fish oils and plasma lipid and lipoprotein metabolism in humans: a critical review. *J Lipid Res*, 1989; **30**:785–807.

[112] Day IN and Wilson DI. Science, medicine, and the future: Genetics and cardiovascular risk. *BMJ*, 2001; **323**:1409–1412.

[113] Prospective Studies Collaboration. Age–specific relevance of usual blood pressure to vascular mortality: a meta–analysis of individual data for one million adults in 61 prospective studies. *Lancet*, 2002; **360**:1903–1913.

[114] Yusuf S, Peto R, Lewis J, Collins R, and Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Progress In Cardiovascular Diseases*, 1985; **27**:335–371.

[115] Blood Pressure Lowering Treatment Trialists' Collaboration. Effects of ACE inhibitors, calcium antagonists, and other blood–pressure–lowering drugs: results of prospectively designed overviews of randomised trials. *Lancet*, 2000; **356**:1955–1964.

[116] Blood Pressure Lowering Trialists' Collaboration. Effects of different blood–pressure–lowering regimens on major cardiovascular events: results of prospectively–designed overviews of randomised trials. *Lancet*, 2003; **362**:1527–1535.

[117] Teo KK, Yusuf S, and Furberg CD. Effects of prophylactic antiarrhythmic drug therapy in acute myocardial infarction. An overview of results from randomized controlled trials. *JAMA*, 1993; **270**:1589–1595.

[118] Heart Outcomes Prevention Evaluation Study Investigators. Effects of an Angiotensin–Converting–Enzyme Inhibitor, Ramipril, on Cardiovascular Events in High–Risk Patients. *N Engl J Med*, 2000; **342**:145–153.

[119] Staessen JA, Fagard R, Thijs L, Celis H, Arabidze GG, Birkenhager WH, Bulpitt CJ, de Leeuw PW, Dollery CT, and Fletcher et al. Randomised double–blind comparison of placebo and active treatment for older patients with isolated systolic hypertension. The Systolic Hypertension in Europe (Syst–Eur) Trial Investigators. *Lancet*, 1997; **350**:757–764.

[120] Staessen JA, Wang JG, and Thijs L. Cardiovascular protection and blood pressure reduction: a meta–analysis. *Lancet*, 2001; **358**:1305–1315.

[121] ALLHAT Collaborative Research Group. Major outcomes in high–risk hypertensive patients randomized to angiotensin–converting enzyme inhibitor or calcium channel blocker vs diuretic: The Antihypertensive and Lipid–Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). *JAMA*, 2002; **288**:2981–2997.

[122] Law MR, Wald NJ, Morris JK, and Jordan RE. Value of low dose combination treatment with blood pressure lowering drugs: analysis of 354 randomised trials. *BMJ*, 2003; **326**:1427–1431.

[123] National High Blood Pressure Education Program Working Group report on primary prevention of hypertension. *Arch Int Med*, 1993; **153**:186–208.

[124] Frost CD, Law MR, and Wald NJ. By how much does dietary salt reduction lower blood pressure? II–Analysis of observational data within populations. *BMJ*, 1991; **302**:815–818.

[125] Elliott P, Stamler J, Nichols R, Dyer AR, Stamler R, Kesteloot H, and Marmot M. Intersalt revisited: further analyses of 24 hour sodium excretion and blood pressure within and across populations. Intersalt Cooperative Research Group. *BMJ*, 1996; **312**:1249–1253.

[126] Law MR, Frost CD, and Wald NJ. By how much does dietary salt reduction lower blood pressure? I–Analysis of observational data among populations. *BMJ*, 1991; **302**:811–815.

[127] Law MR, Frost CD, and Wald NJ. By how much does dietary salt reduction lower blood pressure? III–Analysis of data from trials of salt reduction. *BMJ*, 1991; **302**:819–824.

[128] Neter J, Stam BE, Kok FJ, Grobbee DE, and Geleijnse JM. Influence of Weight Reduction on Blood Pressure: A Meta–Analysis of Randomized Controlled Trials. *Hypertension*, 2003; **42**:878–884.

[129] Appel LJ, Moore TJ, Obarzanek E, Vollmer WM, Svetkey LP, Sacks FM, Bray GA, Vogt TM, Cutler JA, and Windhauser et al. A clinical trial of the effects of dietary patterns on blood pressure. DASH Collaborative Research Group. *N Engl J Med*, 1997; **336**:1117–1124.

[130] Whelton PK, He J, Cutler JA, Brancati FL, Appel LJ, Follmann D, and Klag MJ. Effects of oral potassium on blood pressure. Meta–analysis of randomized controlled clinical trials. *JAMA*, 1997; **277**:1624–1632.

[131] Marmot MG, Elliott P, Shipley MJ, Dyer AR, Ueshima H, Beevers DG, Stamler R, Kesteloot H, Rose G, and Stamler et al. Alcohol and blood pressure: the INTERSALT study. *BMJ*, 1994; **308**:1263–1267.

[132] Barker DJ, Bull AR, Osmond C, and Simmonds SJ. Fetal and placental size and risk of hypertension in adult life. *BMJ*, 1990; **301**:259–262.

[133] Barker DJP, Forsen T, Eriksson JG, and Osmond C. Growth and living conditions in childhood and hypertension in adult life: a longitudinal study. *J Hyperten*, 2002; **20**:1951–1956.

[134] Law CM and Shiell AW. Is blood pressure inversely related to birth weight? The strength of evidence from a systematic review of the literature. *J Hyperten*, 1996; **14**:935–941.

[135] Miura K, Nakagawa H, Tabata M, Morikawa Y, Nishijo M, and Kagamimori S. Birth Weight, Childhood Growth, and Cardiovascular Disease Risk Factors in Japanese Aged 20 Years. *Am J Epidemiol*, 2001; **153**:783–789.

[136] Huxley Rachel, Neil Andrew, and Collins Rory. Unravelling the fetal origins hypothesis: is there really an inverse association between birthweight and subsequent blood pressure? *Lancet*, 2002; **360**:659–665.

[137] Cooper R and Rotimi C. Hypertension in Blacks. *Am J Hyperten*, 1997; **10**:804–812.

[138] Doll R and Hill AB. Smoking and carcinoma of the lung. *BMJ*, 1950; **ii**:739–748.

[139] Wynder EL and Graham EA. Tobacco smoking as a possible etiologic factor in bronchiogenic carcinoma. *JAMA*, 1950; **143**:329–336.

[140] Doll R and Hill AB. The mortality of doctors in relation to their smoking habits. *BMJ*, 1954; **i**:1451–1455.

[141] Hammond EC and Horn D. The relationship between human smoking habits and death rates. *JAMA*, 1954; **155**:1316–1328.

[142] Doll R and Hill AB. Lung cancer and other causes of death in relation to smoking. A second report on the mortality of British doctors. *BMJ*, 1956; **ii**:1071–1076.

[143] Doll R and Hill AB. Mortality in relation to smoking: ten years' observations of British doctors. *BMJ*, 1964; **i**:1399–1414.

[144] Doll R and Peto R. Mortality in relation to smoking: 20 years' observations on male British doctors. *BMJ*, 1976; **ii**:1525–1536.

[145] Doll R, Peto R, Wheatley K, Gray R, and Sutherland I. Mortality in relation to smoking: 40 years' observations on male British doctors. *BMJ*, 1994; **309**:901–911.

[146] Doll R, Peto R, Boreham J, and Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ*, 2004; **328**:1519–0.

[147] Peto R. Smoking and death: the past 40 years and the next 40. *BMJ*, 1994; **309**:937–939.

[148] Ezzati Majid and Lopez Alan D. Estimates of global mortality attributable to smoking in 2000. *Lancet*, 2003; **362**:847–852.

[149] Manson JE, Tosteson H, Ridker PM, Satterfield S, Hebert P, O'Connor GT, Buring JE, and Hennekens CH. The primary prevention of myocardial infarction. *N Engl J Med*, 1992; **326**:1406–1416.

[150] Wilhelmsen L. Coronary heart disease: epidemiology of smoking and intervention studies of smoking. *Am Heart J*, 1988; **115**:242–249.

[151] Parish S, Collins R, Peto R, Youngman L, Barton J, Jayne K, Clarke R, Appleby P, Lyon V, and Cederholm-Williams et al. Cigarette smoking, tar yields, and non–fatal myocardial infarction: 14,000 cases and 32,000 controls in the United Kingdom. The International Studies of Infarct Survival (ISIS) Collaborators. *BMJ*, 1995; **311**:471–477.

[152] Dwyer JH. Exposure to environmental tobacco smoke and coronary risk. *Circulation*, 1997; **96**:1367–1369.

[153] Law MR, Morris JK, and Wald NJ. Environmental tobacco smoke exposure and ischaemic heart disease: an evaluation of the evidence. *BMJ*, 1997; **315**:973–980.

[154] He J, Vupputuri S, Allen K, Prerost MR, Hughes J, and Whelton PK. Passive Smoking and the Risk of Coronary Heart Disease – A meta–analysis of epidemiologic studies. *N Engl J Med*, 1999; **340**:920–926.

[155] Whincup PH, Gilg Julie A, Emberson JR, Jarvis MJ, Feyerabend C, Bryant Andrew, Walker M, and Cook DG. Passive smoking and risk of coronary heart disease and stroke: prospective study using cotinine measurement. *BMJ*, 2004; **329** : 200–205.

[156] Tunstall-Pedoe H, Brown CA, Woodward M, and Tavendale R. Passive smoking by self report and serum cotinine and the prevalence of respiratory and coronary heart disease in the Scottish heart health study. *J Epidemiol Comm Health*, 1995; **49**:139–143.

[157] Meade TW, Imeson J, and Stirling Y. Effects of changes in smoking and other characteristics on clotting factors and the risk of ischaemic heart disease. *Lancet*, 1987; **2**:986–988.

[158] Pittilo RM and Woolf N. Cigarette smoking, endothelial injury and atherosclerosis. *J Smoking–related Dis*, 1993; **4**:17–25.

[159] Hawkins RI. Smoking, platelets and thrombosis. *Nature*, 1972; **236**:450–452.

[160] Cullen P, Schulte H, and Assmann G. Smoking, lipoproteins and coronary heart disease risk: Data from the Munster Heart Study (PROCAM). *Eur Heart J*, 1998; **19**:1632–1641.

[161] Padua MA, Caramelli B, Vianna CB, Chamone D, and Ramires JA. Smoking and lipoprotein abnormalities on platelet aggregation in coronary heart disease. *Int J Cardiol*, 1997; **62**:151–154.

[162] McGill HC. The cardiovascular pathology of smoking. *Am Heart J*, 1988; **115**:250–257.

[163] DHHS (US Department of Health and Human Services). The health benefits of smoking cessation: a report of the Surgeon General. DHHS publication No. (CDC) 90–8416. Washington D.C.: Government Printing Office, 1990.

[164] Peto R, Darby S, Deo H, Silcocks P, Whitley E, and Doll R. Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case–control studies. *BMJ*, 2000; **321**:323–329.

[165] Ockene JK, Kuller LH, Svendsen KH, and Meilahn E. The relationship of smoking cessation to coronary heart disease and lung cancer in the Multiple Risk Factor Intervention Trial (MRFIT). *Am J Public Health*, 1990; **80**:954–958.

[166] Critchley JA and Capewell S. Mortality risk reduction associated with smoking cessation in patients with coronary heart disease: a systematic review. *JAMA*, 2003; **290**:86–97.

[167] McElduff P, Dobson A, Beaglehole R, and Jackson R. Rapid reduction in coronary risk for those who quit cigarette smoking. *ANZ J Public Health*, 1998; **22**:787–791.

[168] Tang JL, Cook DG, and Shaper AG. Giving up smoking: how rapidly does the excess risk of ischaemic heart disease disappear? *J Smoking–related Dis*, 1992; **3**:203–215.

[169] Morris JN, Kagan A, Pattison DC, and Gardner MJ. Incidence and prediction of ischaemic heart–disease in London busmen. *Lancet*, 1966; **2**:553–559.

[170] Morris JN, Everitt MG, Pollard R, Chave SP, and Semmence AM. Vigorous exercise in leisure–time: protection against coronary heart disease. *Lancet*, 1980; **2**:1207–1210.

[171] Morris JN, Clayton DG, Everitt MG, Semmence AM, and Burgess EH. Exercise in leisure time: coronary attack and death rates. *Br Heart J*, 1990; **63**:325–334.

[172] Paffenbarger RS Jr, Hyde RT, Wing AL, and Hsieh CC. Physical activity, all–cause mortality, and longevity of college alumni. *N Engl J Med*, 1986; **314**:605–613.

[173] Berlin JA and Colditz GA. A meta–analysis of physical activity in the prevention of coronary heart disease. *Am J Epidemiol*, 1990; **132**:612–628.

[174] Leon AS, Connett J, Jacobs DR Jr, and Rauramaa R. Leisure–time physical activity levels and risk of coronary heart disease and death. The Multiple Risk Factor Intervention Trial. *JAMA*, 1987; **258**:2388–2395.

[175] Slattery ML, Jacobs DR Jr, and Nichaman MZ. Leisure time physical activity and coronary heart disease death. The US Railroad Study. *Circulation*, 1989; **79**:304–311.

[176] Sesso HD, Paffenbarger RS Jr, and Lee IM. Physical activity and coronary heart disease in men: The Harvard Alumni Health Study. *Circulation*, 2000; **102**:975–980.

[177] Rodriguez BL, Curb JD, Burchfiel CM, Abbott RD, Petrovitch H, Masaki K, and Chiu D. Physical activity and 23–year incidence of coronary heart disease morbidity and mortality among middle–aged men. The Honolulu Heart Program. *Circulation*, 1994; **89**:2540–2544.

[178] Folsom AR, Arnett DK, Hutchinson RG, Liao F, Clegg LX, and Cooper LS. Physical activity and incidence of coronary heart disease in middle–aged women and men. *Med Sci Sports Exerc*, 1997; **29**:901–909.

[179] Kushi LH, Fee RM, Folsom AR, Mink PJ, Anderson KE, and Sellers TA. Physical activity and mortality in postmenopausal women. *JAMA*, 1997; **277**:1287–1292.

[180] Manson JE, Hu FB, Rich-Edwards JW, Colditz GA, Stampfer MJ, Willett WC, Speizer FE, and Hennekens CH. A prospective study of walking as compared with vigorous exercise in the prevention of coronary heart disease in women. *N Engl J Med*, 1999; **341**:650–658.

[181] Sesso HD, Paffenbarger RS, Ha T, and Lee IM. Physical activity and cardiovascular disease risk in middle–aged and older women. *Am J Epidemiol*, 1999; **150**:408–416.

[182] Lee IM, Rexrode KM, Cook NR, Manson JE, and Buring JE. Physical activity and coronary heart disease in women: is 'no pain, no gain' passe? *JAMA*, 2001; **285**:1447–1454.

[183] LaCroix AZ, Leveille SG, Hecht JA, Grothaus LC, and Wagner EH. Does walking decrease the risk of cardiovascular disease hospitalizations and death in older adults? *J Am Geriatr Soc*, 1996; **44**:113–120.

[184] Wannamethee SG, Shaper AG, and Walker M. Changes in physical activity, mortality, and incidence of coronary heart disease in older men. *Lancet*, 1998; **351**:1603–1608.

[185] Bijnen FC, Caspersen CJ, Feskens EJ, Saris WH, Mosterd WL, and Kromhout D. Physical activity and 10–year mortality from cardiovascular diseases and all causes: The Zutphen Elderly Study. *Arch Int Med*, 1998; **158**:1499–1505.

[186] Hakim AA, Curb JD, Petrovitch H, Rodriguez BL, Yano K, Ross GW, White LR, and Abbott RD. Effects of walking on coronary heart disease in elderly men: the Honolulu Heart Program. *Circulation*, 1999; **100**:9–13.

[187] Paffenbarger RS Jr. Physical activity as a defense against coronary heart disease, in Connor WE, Bristow JD (eds): Coronary Heart Disease: prevention, complications and treatment. Philadelphia: JB Lippencott Co, 1985.

[188] Wannamethee SG, Shaper AG, and Walker M. Physical activity and mortality in older men with diagnosed coronary heart disease. *Circulation*, 2000; **102**:1358–1363.

[189] Steffen-Batey L, Nichaman MZ, Goff DC Jr, Frankowski RF, Hanis CL, Ramsey DJ, and Labarthe DR. Change in level of physical activity and risk of all–cause mortality or reinfarction: The Corpus Christi Heart Project. *Circulation*, 2000; **102**:2204–2209.

[190] Hu FB, Stampfer MJ, Solomon C, Liu S, Colditz GA, Speizer FE, Willett WC, and Manson JE. Physical activity and risk for cardiovascular events in diabetic women. *Ann Int Med*, 2001; **134**:96–105.

[191] Lakka TA, Venalainen JM, Rauramaa R, Salonen R, Tuomilehto J, and Salonen JT. Relation of leisure–time physical activity and cardiorespiratory fitness to the risk of acute myocardial infarction. *N Engl J Med*, 1994; **330**:1549–1554.

[192] Mensink GB, Deketh M, Mul MD, Schuit AJ, and Hoffmeister H. Physical activity and its association with cardiovascular risk factors and mortality. *Epidemiol*, 1996; **7**:391–397.

[193] Shaper AG, Wannamethee G, and Weatherall R. Physical activity and ischaemic heart disease in middle–aged British men. *Br Heart J*, 1991; **66**:384–394.

[194] Leon AS and Connett J. Physical activity and 10.5 year mortality in the Multiple Risk Factor Intervention Trial (MRFIT). *Int J Epidemiol*, 1991; **20**:690–697.

[195] Hein HO, Suadicani P, and Gyntelberg F. Physical fitness or physical activity as a predictor of ischaemic heart disease? A 17–year follow–up in the Copenhagen Male Study. *J Int Med*, 1992; **232**:471–479.

[196] Wannamethee SG and Shaper AG. Physical activity in the prevention of cardiovascular disease: an epidemiological perspective. *Sports Med*, 2001; **31**:101–114.

[197] Anderson LB. Physical activity and physical fitness as protection against premature disease or death. *Scand J Med Sci Sports*, 1995; **5**:318–328.

[198] Thune I, Njolstad I, Lochen ML, and Forde OH. Physical activity improves the metabolic risk profiles in men and women: the Tromso Study. *Arch Int Med*, 1998; **158**:1633–1640.

[199] Fung TT, Hu FB, Yu J, Chu NF, Spiegelman D, Tofler GH, Willett WC, and Rimm EB. Leisure–time physical activity, television watching, and plasma biomarkers of obesity and cardiovascular disease risk. *Am J Epidemiol*, 2000; **152**:1171–1178.

[200] Koivisto VA, Yki-Jarvinen H, and DeFronzo RA. Physical training and insulin sensitivity. *Diabetes Metab Rev*, 1986; **1**:445–481.

[201] Henriksson J. Influence of exercise on insulin sensitivity. *J Cardiovasc Risk*, 1995; **2**:303–309.

[202] Wannamethee SG, Shaper AG, and Alberti KG. Physical activity, metabolic factors, and the incidence of coronary heart disease and type 2 diabetes. *Arch Int Med*, 2000; **160**:2108–2116.

[203] Rauramaa R, Salonen JT, Seppanen K, Salonen R, Venalainen JM, Ihanainen M, and Rissanen V. Inhibition of platelet aggregability by moderate–intensity physical exercise: a randomized clinical trial in overweight men. *Circulation*, 1986; **74**:939–944.

[204] Meade TW. Exercise and haemostatic function. *J Cardiovasc Risk*, 1995; **2**:323–329.

[205] Wannamethee S Goya, Lowe Gordon DO, Whincup Peter H, Rumley Ann, Walker Mary, and Lennon Lucy. Physical activity and hemostatic and inflammatory variables in elderly men. *Circulation*, 2002; **105**:1785–1790.

[206] Pyorala K, Laakso M, and Uusitupa M. Diabetes and atherosclerosis: an epidemiologic view. *Diabetes Metab Rev*, 1987; **3**:463–524.

[207] Nathan DM, Meigs J, and Singer DE. The epidemiology of cardiovascular disease in type 2 diabetes mellitus: how sweet it is ... or is it? *Lancet*, 1997; **350**:SI4–SI9.

[208] Klein R. Hyperglycemia and microvascular and macrovascular disease in diabetes. *Diabetes Care*, 1995; **18**:258–268.

[209] Kuusisto J, Mykkanen L, Pyorala K, and Laakso M. NIDDM and its metabolic control predict coronary heart disease in elderly subjects. *Diabetes*, 1994; **43**:960–967.

[210] Lehto S, Ronnemaa T, Haffner SM, Pyorala K, Kallio V, and Laakso M. Dyslipidemia and hyperglycemia predict coronary heart disease events in middle–aged patients with NIDDM. *Diabetes*, 1997; **46**:1354–1359.

[211] Turner RC and Holman RR. The UK Prospective Diabetes Study. UK Prospective Diabetes Study Group. *Ann Med*, 1996; **28**:439–444.

[212] Intensive blood–glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. *Lancet*, 1998; **352**:837–853.

[213] Manson JE, Colditz GA, Stampfer MJ, Willett WC, Krolewski AS, Rosner B, Arky RA, Speizer FE, and Hennekens CH. A prospective study of maturity–onset diabetes mellitus and risk of coronary heart disease and stroke in women. *Arch Int Med*, 1991; **151**:1141–1147.

[214] Stamler J, Vaccaro O, Neaton JD, and Wentworth D. Diabetes, other risk factors, and 12–yr cardiovascular mortality for men screened in the Multiple Risk Factor Intervention Trial. *Diabetes Care*, 1993; **16**:434–444.

[215] Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol–lowering with simvastatin in 5963 people with diabetes: a randomised placebo–controlled trial. *Lancet*, 2003; **361**:2005–2016.

[216] Pyorala K, Pedersen TR, Kjekshus J, Faergeman O, Olsson AG, and Thorgeirsson G. Cholesterol lowering with simvastatin improves prognosis of diabetic patients with coronary heart disease. A subgroup analysis of the Scandinavian Simvastatin Survival Study (4S). *Diabetes Care*, 1997; **20**:614–620.

[217] Keech A, Colquhoun D, Best J, Kirby A, Simes RJ, Hunt D, Hague W, Beller E, Arulchelvam M, and Baker et al. Secondary prevention of cardiovascular events with long–term pravastatin in patients with diabetes or impaired fasting glucose: results from the LIPID trial. *Diabetes Care*, 2003; **26**:2713–2721.

[218] Heart Outcomes Prevention Evaluation Study Investigators. Effects of ramipril on cardiovascular and microvascular outcomes in people with diabetes mellitus: results of the HOPE study and MICRO–HOPE substudy. *Lancet*, 2000; **355**:253–259.

[219] United Kingdom Prospective Diabetes Study Group. Efficacy of atenolol and captopril in reducing risk of macrovascular and microvascular complications in type 2 diabetes. *BMJ*, 1998; **317**:713–720.

[220] United Kingdom Prospective Diabetes Study Group. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes. *BMJ*, 1998; **317**:703–713.

[221] Flegal KM, Carroll MD, Kuczmarski RJ, and Johnson CL. Overweight and obesity in the United States: prevalence and trends, 1960–1994. *Int J Obes Rel Metab Disord*, 1998; **22**:39–47.

[222] Havlik RJ, Hubert HB, Fabsitz RR, and Feinleib M. Weight and hypertension. *Ann Int Med*, 1983; **98**:855–859.

[223] West KM and Kalbfleisch JM. Influence of nutritional factors on prevalence of diabetes. *Diabetes*, 1971; **20**:99–108.

[224] Willett WC, Manson JE, Stampfer MJ, Colditz GA, Rosner B, Speizer FE, and Hennekens CH. Weight, weight change, and coronary heart disease in women. Risk within the 'normal' weight range. *JAMA*, 1995; **273**:461–465.

[225] Spataro JA, Dyer AR, Stamler J, Shekelle RB, Greenlund K, and Garside D. Measures of adiposity and coronary heart disease mortality in the Chicago Western Electric Company Study. *J Clinical Epidemiol*, 1996; **49**:849–857.

[226] Jousilahti P, Tuomilehto J, Vartiainen E, Pekkanen J, and Puska P. Body weight, cardiovascular risk factors, and coronary mortality. 15–year follow–up of middle–aged men and women in eastern Finland. *Circulation*, 1996; **93**:1372–1379.

[227] Shaper AG, Wannamethee SG, and Walker M. Body weight: implications for the prevention of coronary heart disease, stroke, and diabetes mellitus in a cohort study of middle aged men. *BMJ*, 1997; **314**:1311–1317.

[228] Rimm EB, Stampfer MJ, Giovannucci E, Ascherio A, Spiegelman D, Colditz GA, and Willett WC. Body size and fat distribution as predictors of coronary heart disease among middle–aged and older US men. *Am J Epidemiol*, 1995; **141**:1117–1127.

[229] Zhou BF. Effect of body mass index on all–cause mortality and incidence of cardiovascular diseases–report for meta–analysis of prospective studies open optimal cut–off points of body mass index in Chinese adults. *BES*, 2002; **15**:245–252.

[230] Manson JE, Colditz GA, Stampfer MJ, Willett WC, Rosner B, Monson RR, Speizer FE, and Hennekens CH. A prospective study of obesity and risk of coronary heart disease in women. *N Engl J Med*, 1990; **322**:882–889.

[231] Donahue RP, Abbott RD, Bloom E, Reed DM, and Yano K. Central obesity and coronary heart disease in men. *Lancet*, 1987; **1**:821–824.

[232] Folsom AR, Kaye SA, Sellers TA, Hong CP, Cerhan JR, Potter JD, and Prineas RJ. Body fat distribution and 5–year risk of death in older women. *JAMA*, 1993; **269**:483–487.

[233] Prineas RJ, Folsom AR, and Kaye SA. Central adiposity and increased risk of coronary artery disease mortality in older women. *Ann Epidemiol*, 1993; **3**:35–41.

[234] Uwaifo GI and Ratner RE. The roles of insulin resistance, hyperinsulinemia, and thiazolidinediones in cardiovascular disease. *Am J Med*, 2003; **115**:12–19.

[235] Pyorala K, Savolainen E, Kaukola S, and Haapakoski J. Plasma insulin as coronary heart disease risk factor: relationship to other risk factors and predictive value during 9 1/2–year follow–up of the Helsinki Policemen Study population. *Act Med Scand*, 1985; **701**:38–52.

[236] Haffner SM, Valdez RA, Hazuda HP, Mitchell BD, Morales PA, and Stern MP. Prospective analysis of the insulin–resistance syndrome (syndrome X). *Diabetes*, 1992; **41**:715–722.

[237] Zavaroni I, Bonora E, Pagliara M, Dall'Aglio E, Luchetti L, Buonanno G, Bonati PA, Bergonzani M, Gnudi L, and Passeri et al. Risk factors for coronary artery disease in healthy persons with hyperinsulinemia and normal glucose tolerance. *N Engl J Med*, 1989; **320**:702–706.

[238] Despres JP, Lamarche B, Mauriege P, Cantin B, Dagenais GR, Moorjani S, and Lupien PJ. Hyperinsulinemia as an independent risk factor for ischemic heart disease. *N Engl J Med*, 1996; **334**:952–957.

[239] Haffner SM and Miettinen H. Insulin resistance implications for type II diabetes mellitus and coronary heart disease. *Am J Med*, 1997; **103**:152–162.

[240] Hanson RL, Imperatore G, Bennett PH, and Knowler WC. Components of the 'metabolic syndrome' and incidence of type 2 diabetes. *Diabetes*, 2002; **51**:3120–3127.

[241] Lakka HM, Laaksonen DE, Lakka TA, Niskanen LK, Kumpusalo E, Tuomilehto J, and Salonen JT. The metabolic syndrome and total and cardiovascular disease mortality in middle–aged men. *JAMA*, 2002; **288**:2709–2716.

[242] Golden SH, Folsom AR, Coresh J, Sharrett AR, Szklo M, and Brancati F. Risk factor groupings related to insulin resistance and their synergistic effects on subclinical atherosclerosis: the atherosclerosis risk in communities study. *Diabetes*, 2002; **51**:3069–3076.

[243] Girman CJ, Rhodes T, Mercuri M, Pyorala K, Kjekshus J, Pedersen TR, Beere PA, Gotto AM, and Clearfield M. The metabolic syndrome and risk of major coronary events in the Scandinavian Simvastatin Survival Study (4S) and the Air Force/Texas Coronary Atherosclerosis Prevention Study (AFCAPS/TexCAPS). *Am J Cardiol*, 2004; **93**:136–141.

[244] Rich-Edwards JW, Manson JE, Hennekens CH, and Buring JE. The primary prevention of coronary heart disease in women. *N Engl J Med*, 1995; **26**:1758-1766.

[245] Barrett-Connor E. Sex differences in coronary heart disease. Why are women so superior? The 1995 Ancel Keys lecture. *Circulation*, 1997; **95**:252-264.

[246] Waldron I. Why do women live longer than men? *Soc Sci Med*, 1976; **10**:349-362.

[247] Wingard DL, Suarez L, and Barrett-Connor E. The sex differential in mortality from all causes and ischaemic heart disease. *Am J Epidemiol*, 1983; **117**:165-172.

[248] Stampfer MJ and Colditz GA. Estrogen replacement therapy and coronary heart disease: a quantitative assessment of the epidemiologic evidence. *Prev Med*, 1991; **20**:47-63.

[249] Hulley S, Grady D, Furberg C, Herrington D, Riggs B, and Vittinghoff E. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. Heart and Estrogen/Progestin Replacement Study (HERS) research group. *JAMA*, 1998; **280**:605-613.

[250] Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, Jackson RD, Beresford SA, Howard BV, Johnson KC, Kotchen JM, and Ockene J. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principle results from the Women's Health Initiative (WHI) randomised controlled trial. *JAMA*, 2002; **288**:321-333.

[251] The Women's Health Initiative Steering Committee. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative (WHI) randomised controlled trial. *JAMA*, 2004; **291**:1701-1712.

[252] Lawlor DA, Davey Smith G, and Ebrahim S. Socioeconomic position and hormone replacement therapy use: explaining the discrepancy in evidence from observational and randomized controlled trials. *Am J Public Health*, 2004; **94**:2149-2154.

[253] Petitti D. Commentary: hormone replacement therapy and coronary heart disease: four lessons. *Int J Epidemiol*, 2004; **33**:461-463.

[254] Lawlor DA, Davey Smith G, and Ebrahim S. Commentary: the hormone replacement–coronary heart disease conundrum. *Int J Epidemiol*, 2004; **33**:464-467.

[255] Beilin LJ and Puddey IB. Alcohol and cardiovascular disease–more than one paradox to consider. *J Cardiovasc Risk*, 2003; **10**:1–30.

[256] Maclure M. Demonstration of deductive meta–analysis: ethanol intake and risk of myocardial infarction. *Epidemiol Rev*, 1993; **15**:328–351.

[257] Corrao G, Bagnardi V, Zambon A, and La Vecchia C. A meta–analysis of alcohol consumption and the risk of 15 diseases. *Prev Med*, 2004; **38**:613–619.

[258] Thun MJ, Peto R, Lopez AD, Monaco JH, Henley SJ, Heath CW Jr, and Doll R. Alcohol consumption and mortality among middle–aged and elderly U.S. adults. *N Engl J Med*, 1997; **337**:1705–1714.

[259] Fuchs CS, Stampfer MJ, Colditz GA, Giovannucci EL, Manson JE, Kawachi I, Hunter DJ, Hankinson SE, Hennekens CH, and Rosner et al. Alcohol consumption and mortality among women. *N Engl J Med*, 1995; **332**:1245–1250.

[260] Doll R, Peto R, Hall E, Wheatley K, and Gray R. Mortality in relation to consumption of alcohol: 13 years' observations on male British doctors. *BMJ*, 1994; **309**:911–918.

[261] Farchi G, Fidanza F, Mariotti S, and Menotti A. Alcohol and mortality in the Italian rural cohorts of the Seven Countries Study. *Int J Epidemiol*, 1992; **21**:74–81.

[262] Klatsky AL, Armstrong MA, and Friedman GD. Alcohol and mortality. *Ann Int Med*, 1992; **117**:646–654.

[263] Britton A and Marmot M. Different measures of alcohol consumption and risk of coronary heart disease and all–cause mortality: 11–year follow–up of the Whitehall II Cohort Study. *Addiction*, 2004; **99**:109–116.

[264] Gaziano JM, Gaziano TA, Glynn RJ, Sesso HD, Ajani UA, Stampfer MJ, Manson JE, Hennekens CH, and Buring JE. Light–to–moderate alcohol consumption and mortality in the physicians health study enrollment cohort. *J Am Coll Cardiol*, 2000; **35**:96–105.

[265] Shaper AG, Wannamethee G, and Walker M. Alcohol and mortality in British men: explaining the U–shaped curve. *Lancet*, 1988; **2**:1267–1273.

[266] Mukamal KJ, Conigrave KM, Mittleman MA, Camargo CA Jr, Stampfer MJ, Willett WC, and Rimm EB. Roles of drinking pattern and type of alcohol consumed in coronary heart disease in men. *N Engl J Med*, 2003; **348**:109–118.

[267] Rimm EB, Giovannucci EL, Willett WC, Colditz GA, Ascherio A, Rosner B, and Stampfer MJ. Prospective study of alcohol consumption and risk of coronary disease in men. *Lancet*, 1991; **338**:464–468.

[268] Rimm EB, Williams P, Fosher K, Criqui M, and Stampfer MJ. Moderate alcohol intake and lower risk of coronary heart disease: meta–analysis of effects on lipids and haemostatic factors. *BMJ*, 1999; **319**:1523–1528.

[269] Langer RD, Criqui MH, and Reed DM. Lipoproteins and blood pressure as biological pathways for effect of moderate alcohol consumption on coronary heart disease. *Circulation*, 1992; **85**:910–915.

[270] Suh I, Shaten BJ, Cutler JA, and Kuller LH. Alcohol use and mortality from coronary heart disease: the role of high–density lipoprotein cholesterol. The Multiple Risk Factor Intervention Trial Research Group. *Ann Int Med*, 1992; **116**:881–887.

[271] Gaziano JM, Buring JE, Breslow JL, Goldhaber SZ, Rosner B, VanDenburgh M, Willett W, and Hennekens CH. Moderate alcohol intake, increased levels of high–density lipoprotein and its subfractions, and decreased risk of myocardial infarction. *N Engl J Med*, 1993; **329**:1829–1834.

[272] Marques-Vidal P, Ducimetiere P, Evans A, Cambou JP, and Arveiler D. Alcohol consumption and myocardial infarction: a case–control study in France and Northern Ireland. *Am J Epidemiol*, 1996; **143**:1089–1093.

[273] Renaud SC, Beswick AD, Fehily AM, Sharp DS, and Elwood PC. Alcohol and platelet aggregation: the Caerphilly Prospective Heart Disease Study. *Am J Clin Nutr*, 1992; **55**:1012–1017.

[274] Renaud S and de Lorgeril M. Wine, alcohol, platelets, and the French paradox for coronary heart disease. *Lancet*, 1992; **339**:1523–1526.

[275] Folsom AR, Wu KK, Davis CE, Conlan MG, Sorlie PD, and Szklo M. Population correlates of plasma fibrinogen and factor VII, putative cardiovascular risk factors. *Atherosclerosis*, 1991; **91**:191–205.

[276] Meade TW, Chakrabarti R, Haines AP, North WR, and Stirling Y. Characteristics affecting fibrinolytic activity and plasma fibrinogen concentrations. *BMJ*, 1979; **1**:153–156.

[277] Gronbaek M, Becker U, Johansen D, Gottschau A, Schnohr P, Hein HO, Jensen G, and Sorensen TIA. Type of Alcohol Consumed and Mortality from All Causes, Coronary Heart Disease, and Cancer. *Ann Int Med*, 2000; **133**:411–419.

[278] Wannamethee SG and Shaper AG. Type of alcoholic drink and risk of major coronary heart disease events and all–cause mortality. *Am J Public Health*, 1999; **89**:685–690.

[279] Theobald H, Bygren LO, Carstensen J, and Engfeldt P. A moderate intake of wine is associated with reduced total mortality and reduced mortality from cardiovascular disease. *J Stud Alcohol*, 2000; **61**:652–656.

[280] Di Castelnuovo A, Rotondo S, Iacoviello L, Donati MB, and De Gaetano G. Meta–analysis of wine and beer consumption in relation to vascular risk. *Circulation*, 2002; **105**:2836–2844.

[281] Iijima K, Yoshizumi M, and Ouchi Y. Effect of red wine polyphenols on vascular smooth muscle cell function–molecular mechanism of the 'French paradox'. *Mech Ageing Dev*, 2002; **123**:1033–1039.

[282] Mar MH and Zeisel SH. Betaine in wine: answer to the French paradox? *Med Hypotheses*, 1999; **53**:383–385.

[283] Inhibition of LDL oxidation by phenolic substances in red wine: a clue to the French paradox? *Nutr Rev*, 1993; **51**:185–187.

[284] Rehm J, Sempos CT, and Trevisan M. Alcohol and cardiovascular disease–more than one paradox to consider. Average volume of alcohol consumption, patterns of drinking and risk of coronary heart disease–a review. *J Cardiovasc Risk*, 2003; **10**:15–20.

[285] Gruchow HW, Hoffmann RG, Anderson AJ, and Barboriak JJ. Effects of drinking patterns on the relationship between alcohol an coronary occlusion. *Atherosclerosis*, 1982; **43**:393–404.

[286] Kauhanen J, Kaplan GA, Goldberg DE, and Salonen JT. Beer binging and mortality: results from the Kuopio ischaemic heart disease risk factor study, a prospective population based study. *BMJ*, 1997; **315**:846–851.

[287] McElduff P and Dobson AJ. How much alcohol and how often? Population based case–control study of alcohol consumption and risk of a major coronary event. *BMJ*, 1997; **314**:1159–1164.

[288] Rehm J, Greenfield TK, and Rogers JD. Average volume of alcohol consumption, patterns of drinking, and all–cause mortality: results from the US National Alcohol Survey. *Am J Epidemiol*, 2001; **153**:64–71.

[289] Rakic V, Puddey IB, Dimmitt SB, Burke V, and Beilin LJ. A controlled trial of the effects of pattern of alcohol intake on serum lipid levels in regular drinkers. *Atherosclerosis*, 1998; **137**:243–252.

[290] Wannamethee G and Shaper AG. Alcohol intake and variations in blood pressure by day of examination. *J Hum Hypertens*, 1991; **5**:59–67.

[291] Rakic V, Puddey IB, Burke V, Dimmitt SB, and Beilin LJ. Influence of pattern of alcohol intake on blood pressure in regular drinkers: a controlled trial. *J Hyperten*, 1998; **16**:165–174.

[292] Abe H, Kawano Y, Kojima S, Ashida T, Kuramochi M, Matsuoka H, and Omae T. Biphasic effects of repeated alcohol intake on 24–hour blood pressure in hypertensive patients. *Circulation*, 1994; **89**:2626–2633.

[293] Mikhailidis DP, Barradas MA, and Jeremy JY. The effect of ethanol on platelet function and vascular prostanoids. *Alcohol*, 1990; **7**:171–180.

[294] Selhub J, Jacques PF, Wilson PW, Rush D, and Rosenberg IH. Vitamin status and intake as primary determinants of homocysteinemia in an elderly population. *JAMA*, 1993; **270**:2693–2698.

[295] Clarke R, Woodhouse P, Ulvik A, Frost C, Sherliker P, Refsum H, Ueland PM, and Khaw KT. Variability and determinants of total homocysteine concentrations in plasma in an elderly population. *Clin Chem*, 1998; **44**:102–107.

[296] Mudd SH, Skovby F, Levy HL, Pettigrew KD, Wilcken B, and Pyeritz RE. The natural history of homocystinuria due to cystathionine β–synthase deficiency. *Am J Hum Genet*, 1997; **37**:1–31.

[297] McCully KS. Vascular pathology of homocysteinemia: implications for the pathogenesis of arteriosclerosis. *Am J Pathol*, 1969; **56**:111–128.

[298] Boers GH, Smals AG, Trijbels FJ, Fowler B, Bakkeren JA, Schoonderwaldt HC, Kleijer WJ, and Kloppenborg PW. Heterozygosity for homocystinuria in premature peripheral and cerebral occlusive arterial disease. *N Engl J Med*, 1985; **313**:709–715.

[299] Clarke R, Daly L, Robinson K, Naughten E, Cahalane S, Fowler B, and Graham I. Hyperhomocysteinemia: an independent risk factor for vascular disease. *N Engl J Med*, 1991; **324**:1149–1155.

[300] Homocysteine Studies Collaboration. Homocysteine and risk of ischemic heart disease and stroke: a meta–analysis. *JAMA*, 2002; **288**:2015–2022.

[301] Wald DS, Law M, and Morris JK. Homocysteine and cardiovascular disease: evidence on causality from a meta–analysis. *BMJ*, 2002; **325**:1202–1206.

[302] Collaboration Homocysteine Lowering Trialists'. Lowering blood homocysteine with folic acid based supplements: meta–analysis of randomised trials. *BMJ*, 1998; **316**:894–898.

[303] Jacques PF, Selhub J, Bostom AG, Wilson PWF, and Rosenberg IH. The Effect of Folic Acid Fortification on Plasma Folate and Total Homocysteine Concentrations. *N Engl J Med*, 1999; **340**:1449–1454.

[304] Clarke R and Collins R. Can dietary supplements with folic acid or vitamin B6 reduce cardiovascular risk? Design of clinical trials to test the homocysteine hypothesis of vascular disease. *J Cardiovasc Risk*, 1998; **4**:249-255.

[305] Baker F and Picton D. Blinded comparison of folic acid and placebo in patients with ischaemic heart disease: an outcome trial. *Circulation*, 2002; A3642.

[306] Toole JF, Malinow MR, Chambless LE, Spence JD, Pettigrew LC, Howard VJ, Sides EG, Wang CH, and Stampfer M. Lowering homocysteine in patients with ischaemic stroke to prevent recurrent stroke, myocardial infarction, and death: the Vitamin Intervention for Stroke Prevention (VISP) randomized controlled trial. *JAMA*, 2004; **291**:565-575.

[307] Harmon DL, Woodside JV, Yarnell JW, McMaster D, Young IS, McCrum EE, Gey KF, Whitehead AS, and Evans AE. The common 'thermolabile' variant of methylene tetrahydrofolate reductase is a major determinant of mild hyperhomocysteinaemia. *Q J Med*, 1996; **89**:571–577.

[308] Jacques PF, Bostom AG, Williams RR, Ellison RC, Eckfeldt JH, Rosenberg IH, Selhub J, and Rozen R. Relation between folate status, a common mutation in methylenetetrahydrofolate reductase, and plasma homocysteine concentrations. *Circulation*, 1996; **93**:7–9.

[309] Oliver MF. Antioxidant nutrients, atherosclerosis, and coronary heart disease. *Br Heart J*, 1995; **73**:299–301.

[310] Gale CR, Martyn CN, Winter PD, and Cooper C. Vitamin C and risk of death from stroke and coronary heart disease in cohort of elderly people. *BMJ*, 1995; **310**:1563–1566.

[311] Enstrom JE, Kanim LE, and Klein MA. Vitamin C intake and mortality among a sample of the United States population. *Epidemiol*, 1992; **3**:194–202.

[312] Gey KF, Stahelin HB, and Eichholzer M. Poor plasma status of carotene and vitamin C is associated with higher mortality from ischemic heart disease and stroke: Basel Prospective Study. *Clinical Invest*, 1993; **71**:3–6.

[313] Khaw KT, Bingham S, Welch A, Luben R, Wareham N, Oakes S, and Day N. Relation between plasma ascorbic acid and mortality in men and women in EPIC–Norfolk prospective study: a prospective population study. European Prospective Investigation into Cancer and Nutrition. *Lancet*, 2001; **357**:657–663.

[314] Knekt P, Reunanen A, Jarvinen R, Seppanen R, Heliovaara M, and Aromaa A. Antioxidant vitamin intake and coronary mortality in a longitudinal population study. *Am J Epidemiol*, 1994; **139**:1180–1189.

[315] Losonczy KG, Harris TB, and Havlik RJ. Vitamin E and vitamin C supplement use and risk of all–cause and coronary heart disease mortality in older persons: the Established Populations for Epidemiologic Studies of the Elderly. *Am J Clin Nutr*, 1996; **64**:190–196.

[316] Pandey DK, Shekelle R, Selwyn BJ, Tangney C, and Stamler J. Dietary vitamin C and beta–carotene and risk of death in middle–aged men. The Western Electric Study. *Am J Epidemiol*, 1995; **142**:1269–1278.

[317] Rimm EB, Stampfer MJ, Ascherio A, Giovannucci E, Colditz GA, and Willett WC. Vitamin E consumption and the risk of coronary heart disease in men. *N Engl J Med*, 1993; **328**:1450–1456.

[318] Stampfer MJ, Hennekens CH, Manson JE, Colditz GA, Rosner B, and Willett WC. Vitamin E consumption and the risk of coronary disease in women. *N Engl J Med*, 1993; **328**:1444–1449.

[319] Asplund K. Antioxidant vitamins in the prevention of cardiovascular disease: a systematic review. *J Intern Med*, 2002; **251**:372–392.

[320] Pereira MA, O'Reilly E, Augustsson K, Fraser GE, Goldbourt U, Heitmann BL, Hallmans G, Knekt P, Liu S, and Pietinen P. Dietary Fiber and Risk of Coronary Heart Disease: A Pooled Analysis of Cohort Studies. *Arch Int Med*, 2004; **164**:370–376.

[321] Dietary supplementation with n–3 polyunsaturated fatty acids and vitamin E after myocardial infarction: results of the GISSI–Prevenzione trial. . *Lancet*, 1999; **354**:447–455.

[322] Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of antioxidant vitamin supplementation in 20,536 high–risk individuals: a randomised placebo–controlled trial. *Lancet*, 2002; **360**:23–33.

[323] Rapola JM, Virtamo J, Ripatti S, Huttunen JK, Albanes D, Taylor PR, and Heinonen OP. Randomised trial of alpha–tocopherol and beta–carotene supplements on incidence of major coronary events in men with previous myocardial infarction. *Lancet*, 1997; **349**:1715–1720.

[324] Stephens NG, Parsons A, Schofield PM, Kelly F, Cheeseman K, and Mitchinson MJ. Randomised controlled trial of vitamin E in patients with coronary disease: Cambridge Heart Antioxidant Study (CHAOS). *Lancet*, 1996; **347**:781–786.

[325] Heart Outcomes Prevention Evaluation Study Investigators. Vitamin E supplementation and cardiovascular events in high–risk patients. . *N Engl J Med*, 2000; **342**:154–160.

[326] Hennekens CH, Buring JE, Manson JE, Stampfer M, Rosner B, Cook NR, Belanger C, LaMotte F, Gaziano JM, and Ridker et al. Lack of effect of long–term supplementation with beta carotene on the incidence of malignant neoplasms and cardiovascular disease. *N Engl J Med*, 1996; **334**:1145–1149.

[327] Virtamo J, Rapola JM, Ripatti S, Heinonen OP, Taylor PR, Albanes D, and Huttunen JK. Effect of vitamin E and beta carotene on the incidence of primary nonfatal myocardial infarction and fatal coronary heart disease. *Arch Int Med*, 1998; **158**:668–675.

[328] Greenberg ER, Baron JA, Karagas MR, Stukel TA, Nierenberg DW, Stevens MM, Mandel JS, and Haile RW. Mortality associated with low plasma concentration of beta carotene and the effect of oral supplementation. *JAMA*, 1996; **275**:699–703.

[329] Collaborative Group of the Primary Prevention Project. Low–dose aspirin and vitamin E in people at cardiovascular risk: a randomised trial in general practice. *Lancet*, 2001; **357**:89–95.

[330] Blot WJ, Li JY, Taylor PR, Guo W, Dawsey S, Wang GQ, Yang CS, Zheng SF, Gail M, and Li et al. Nutrition intervention trials in Linxian, China: supplementation with specific vitamin/mineral combinations, cancer incidence, and disease–specific mortality in the general population. *J Nat Canc Inst*, 1993; **85**:1483–1492.

[331] Lee IM, Cook NR, Manson JE, Buring JE, and Hennekens CH. Beta–carotene supplementation and incidence of cancer and cardiovascular disease: the Women's Health Study. *J Nat Can Instit*, 1999; **91**:2102–2106.

[332] Morris CD and Carson S. Routine vitamin supplementation to prevent cardiovascular disease: a summary of the evidence for the U.S. Preventive services task force. *Ann Int Med*, 2003; **139**:56–70.

[333] Steinberg D. Clinical trials of antioxidants in atherosclerosis: are we doing the right thing? *Lancet*, 1995; **346**:36–38.

[334] Ross R. Atherosclerosis–an inflammatory disease. *N Engl J Med*, 1999; **340**:115–126.

[335] Libby P, Egan D, and Skarlatos S. Roles of Infectious Agents in Atherosclerosis and Restenosis: An Assessment of the Evidence and Need for Future Research. *Circulation*, 1997; **96**:4095–4103.

[336] Saikku P, Leinonen M, Mattila K, Ekman MR, Nieminen MS, Makela PH, Huttunen JK, and Valtonen V. Serological evidence of an association of a novel Chlamydia, TWAR, with chronic coronary heart disease and acute myocardial infarction. *Lancet*, 1988; **2**:983–986.

[337] Danesh J, Collins R, and Peto R. Chronic infections and coronary heart disease: is there a link? *Lancet*, 1997; **350**:430–436.

[338] Danesh J and Appleby P. Persistant infection and vascular disease: a systematic review. *Expert Opin Invest Drugs*, 1998; **7**:691–713.

[339] Wong YK, Gallagher PJ, and Ward ME. Chlamydia pneumoniae and atherosclerosis. *Heart*, 1999; **81**:232–238.

[340] Danesh J, Whincup P, Walker M, Lennon L, Thomson A, Appleby P, Wong Y, Bernardes-Silva M, and Ward M. Chlamydia pneumoniae IgG titres and coronary heart disease: prospective study and meta–analysis. *BMJ*, 2000; **321**:208–213.

[341] Danesh J. Coronary heart disease, Helicobacter pylori, dental disease, Chlamydia pneumoniae, and cytomegalovirus: meta–analyses of prospective studies. *Am Heart J*, 1999; **138**:S434–S437.

[342] Whincup P, Danesh J, Walker M, Lennon L, Thomson A, Appleby P, Hawkey C, and Atherton J. Prospective study of potentially virulent strains of Helicobacter pylori and coronary heart disease in middle–aged men. *Circulation*, 2000; **101**:1647–1652.

[343] Danesh J, Collins R, Appleby P, and Peto R. Association of fibrinogen, C–reactive protein, albumin, or leukocyte count with coronary heart disease: meta–analyses of prospective studies. *JAMA*, 1998; **279**:1477–1482.

[344] Ridker PM, Buring JE, Cook NR, and Rifai N. C–reactive protein, the metabolic syndrome, and risk of incident cardiovascular events: an 8–year follow–up of 14 719 initially healthy American women. *Circulation*, 2003; **107**:391–397.

[345] Kuller LH, Tracy RP, Shaten J, and Meilahn EN. Relation of C–reactive protein and coronary heart disease in the MRFIT nested case–control study. Multiple Risk Factor Intervention Trial. *Am J Epidemiol*, 1996; **144**:537–547.

[346] Danesh J, Wheeler JG, Hirschfield GM, Eda S, Eiriksdottir G, Rumley A, Lowe GDO, Pepys M B, and Gudnason V. C–Reactive Protein and Other Circulating Markers of Inflammation in the Prediction of Coronary Heart Disease. *N Engl J Med*, 2004; **350**:1387–1397.

[347] O'Connor CM, Dunne MW, Pfeffer MA, Muhlestein JB, Yao L, Gupta S, Benner RJ, Fisher MR, and Cook TD. Azithromycin for the secondary prevention of coronary heart disease events the WIZARD Study: a randomized controlled trial. *JAMA*, 2003; **290**:1459-1466.

[348] Antibiotic Treatment for Secondary Prevention of Coronary Events, Results of Azithromycin and Coronary Events Study (ACES), European Society of Cardiology, Munich, 2004. Available at http://www.escardio.org/nr/rdonlyres/4a72d326-1d39-4243-8689-03fe9d423d8e/0/aceshotlineiii.ppt. Accessed December 10, 2004.

[349] PROVE-IT TIMI 22 Antibiotic Trial, European Society of Cardiology, Munich, 2004. Avaiable at http://www.escardio.org/nr/rdonlyres/ca014c64-5a46-4489-a30b-b43141ddbe0f/0/prove_ittimi_22hotlineiii.ppt. Accessed December 10, 2004.

[350] Thaulow E, Erikssen J, Sandvik L, Stormorken H, and Cohn PF. Blood platelet count and function are related to total and cardiovascular death in apparently healthy men. *Circulation*, 1991; **84**:613–617.

[351] Elwood PC, Renaud S, Sharp DS, Beswick AD, O'Brien JR, and Yarnell JW. Ischemic heart disease and platelet aggregation. The Caerphilly Collaborative Heart Disease Study. *Circulation*, 1991; **83**:38–44.

[352] Junker R, Heinrich J, Schulte H, van de Loo J, and Assmann G. Coagulation factor VII and the risk of coronary heart disease in healthy men. *Arteriosclerosis, Thrombosis, And Vascular Biology*, 1997; **17**:1539–1544.

[353] Wilhelmsen L, Svardsudd K, Korsan-Bengtsen K, Larsson B, Welin L, and Tibblin G. Fibrinogen as a risk factor for stroke and myocardial infarction. *N Engl J Med*, 1984; **311**:501–505.

[354] Meade TW, Mellows S, Brozovic M, Miller GJ, Chakrabarti RR, North WR, Haines AP, Stirling Y, Imeson JD, and Thompson et al. Haemostatic function and ischaemic heart disease: principal results of the Northwick Park Heart Study. *Lancet*, 1986; **2**:533–537.

[355] Woodward M, Lowe GDO, Rumley A, and Tunstall-Pedoe H. Fibrinogen as a risk factor for coronary heart disease and mortality in middle–aged men and women: The Scottish Heart Health Study. *Eur Heart J*, 1998; **19**:55–62.

[356] Danesh J, Whincup P, Walker M, Lennon L, Thomson A, Appleby P, Rumley A, and Lowe GD. Fibrin D–dimer and coronary heart disease: prospective study and meta–analysis. *Circulation*, 2001; **103**:2323–2327.

[357] Lowe GDO, Danesh J, Lewington S, Walker M, Lennon L, Thomson A, Rumley A, and Whincup PH. Tissue plasminogen activator antigen and coronary heart disease: prospective study and meta–analysis. *Eur Heart J*, 2004; **25**:252-259.

[358] Whincup P, Danesh J, Walker M, Lennon L, Thomson A, Appleby P, Rumley A, and Lowe GDO. von Willebrand factor and coronary heart disease: prospective study and meta–analysis. *Eur Heart J*, 2002; **23**:1764–1770.

[359] Antithrombotic Trialists' Collaboration. Collaborative meta–analysis of randomised trials of antiplatelet therapy for prevention of death, myocardial infarction, and stroke in high risk patients. *BMJ*, 2002; **324**:71–86.

[360] Fibrinolytic Therapy Trialists' (FTT) Collaborative Group. Indications for fibrinolytic therapy in suspected acute myocardial infarction: collaborative overview of early mortality and major morbidity results from all randomised trials of more than 1000 patients. *Lancet*, 1994; **343**:311–322.

[361] Marmot MG, Smith GD, Stansfeld S, Patel C, North F, Head J, White I, Brunner E, and Feeney A. Health inequalities among British civil servants: the Whitehall II study. *Lancet*, 1991; **337**:1387–1393.

[362] Marmot MG, Rose G, Shipley M, and Hamilton PJ. Employment grade and coronary heart disease in British civil servants. *J Epidemiol Comm Health*, 1978; **32**:244–249.

[363] Marmot MG, Bosma H, Hemingway H, Brunner E, and Stansfeld S. Contribution of job control and other risk factors to social variations in coronary heart disease incidence. *Lancet*, 1997; **350**:235–239.

[364] Bosma H, Marmot MG, Hemingway H, Nicholson AC, Brunner E, and Stansfeld SA. Low job control and risk of coronary heart disease in Whitehall II (prospective cohort) study. *BMJ*, 1997; **314**:558–565.

[365] Michie S and Cockcroft A. Overwork can kill. *BMJ*, 1996; **312**:921–922.

[366] Barefoot JC and Schroll M. Symptoms of depression, acute myocardial infarction, and total mortality in a community sample. *Circulation*, 1996; **93**:1976–1980.

[367] Anda R, Williamson D, Jones D, Macera C, Eaker E, Glassman A, and Marks J. Depressed affect, hopelessness, and the risk of ischemic heart disease in a cohort of U.S. adults. *Epidemiol*, 1993; **4**:285–294.

[368] Hemingway H and Marmot M. Psychosocial factors in the aetiology and prognosis of coronary heart disease. Systematic review of prospective cohort studies. *BMJ*, 1999; **318**:1460–1467.

[369] Williams RB Jr, Haney TL, Lee KL, Kong YH, Blumenthal JA, and Whalen RE. Type A behavior, hostility, and coronary atherosclerosis. *Psychosom Med*, 1980; **42**:539–549.

[370] Barefoot JC, Dahlstrom WG, and Williams RB Jr. Hostility, CHD incidence, and total mortality: a 25–year follow–up study of 255 physicians. *Psychosomc Med*, 1983; **45**:59–63.

[371] Brunner E, Davey Smith G, Marmot M, Canner R, Beksinska M, and O'Brien J. Childhood social circumstances and psychosocial and behavioural factors as determinants of plasma fibrinogen. *Lancet*, 1996; **347**:1008–1013.

[372] Steptoe A, Kunz-Ebrecht S, Owen N, Feldman PJ, Rumley A, Lowe GDO, and Marmot M. Influence of socioeconomic status and job control on plasma fibrinogen responses to acute mental stress. *Psych Med*, ; **65**:137–144.

[373] Rosengren A, Wilhelmsen L, Welin L, Tsipogianni A, Teger-Nilsson AC, and Wedel H. Social influences and cardiovascular risk factors as determinants of plasma fibrinogen concentration in a general population sample of middle aged men. *BMJ*, 1990; **300**:634–638.

[374] Carney RM, Saunders RD, Freedland KE, Stein P, Rich MW, and Jaffe AS. Association of depression with reduced heart rate variability in coronary artery disease. *Am J Cardiol*, 1995; **76**:562–564.

[375] McCraty R, Atkinson M, Tiller WA, Rein G, and Watkins AD. The effects of emotions on short–term power spectrum analysis of heart rate variability. *Am J Cardiol*, 1995; **76**:1089–1093.

[376] Hemingway H and Marmot M. Psychosocial factors in the primary and secondary prevention of coronary heart disease: a systematic review. In: Yusuf S, Cairns JA, Gersch BJ, Camm AJ (eds.) Evidence Based Cardiology. London: British Medical Journal Publishing, 1997.

[377] Barker DJ, Gluckman PD, Godfrey KM, Harding JE, Owens JA, and Robinson JS. Fetal nutrition and cardiovascular disease in adult life. *Lancet*, 1993; **341**:938–941.

[378] Barker DJP. Fetal and infant origins of adult disease. London: British Medical Journal, 1992.

[379] Paneth N and Susser M. Early origin of coronary heart disease (the 'Barker hypothesis'). *BMJ*, 1995; **310**:411–412.

[380] Barker DJ, Osmond C, Simmonds SJ, and Wield GA. The relation of small head circumference and thinness at birth to death from cardiovascular disease in adult life. *BMJ*, 1993; **306**:422–426.

[381] Barker DJ, Winter PD, Osmond C, Margetts B, and Simmonds SJ. Weight in infancy and death from ischaemic heart disease. *Lancet*, 1989; **2**:577–580.

[382] Leon DA, Lithell HO, Vagero D, Koupilova I, Mohsen R, Berglund L, Lithell UB, and McKeigue PM. Reduced fetal growth rate and increased risk of death from ischaemic heart disease: cohort study of 15 000 Swedish men and women born 1915–29. *BMJ*, 1998; **317**:241–245.

[383] Rich-Edwards JW, Stampfer MJ, Manson JE, Rosner B, Hankinson SE, Colditz GA, Willett WC, and Hennekens CH. Birth weight and risk of cardiovascular disease in a cohort of women followed up since 1976. *BMJ*, 1997; **315**:396–400.

[384] Barker DJP, Martyn CN, Osmond C, Hales CN, and Fall CHD. Growth in utero and serum cholesterol concentrations in adult life. *BMJ*, 1993; **307**:1524–1527.

[385] Martyn CN, Gale CR, Jespersen S, and Sherriff SB. Impaired fetal growth and atherosclerosis of carotid and peripheral arteries. *Lancet*, 1998; **352**:173–178.

[386] Suzuki T, Minami J, Ohrui M, Ishimitsu T, and Matsuoka H. Relationship between birth weight and cardiovascular risk factors in Japanese young adults. *Am J Hyperten*, 2000; **13**:907–913.

[387] Hales CN, Barker DJ, Clark PM, Cox LJ, Fall C, Osmond C, and Winter PD. Fetal and infant growth and impaired glucose tolerance at age 64. *BMJ*, 1991; **303**:1019–1022.

[388] Barker DJP, Meade TW, Fall CHD, Lee A, Osmond C, Phipps K, and Stirling Y. Relation of fetal and infant growth to plasma fibrinogen and factor VII concentrations in adult life. *BMJ*, 1992; **304**:148–152.

[389] Owen CG, Whincup PH, Odoki K, Gilg JA, and Cook DG. Birth Weight and Blood Cholesterol Level: A Study in Adolescents and Systematic Review. *Pediatrics*, 2003; **111**:1081–1089.

[390] Barker DJ, Osmond C, and Golding J. Height and mortality in the counties of England and Wales. *Annals of Human Biology*, 1990; **17**:1–6.

[391] Hebert PR, Rich-Edwards JW, Manson JE, Ridker PM, Cook NR, O'Connor GT, Buring JE, and Hennekens CH. Height and incidence of cardiovascular disease in male physicians. *Circulation*, 1993; **88**:1437–1443.

[392] Rich-Edwards JW, Manson JE, Stampfer MJ, Colditz GA, Willett WC, Rosner B, Speizer FE, and Hennekens CH. Height and the risk of cardiovascular disease in women. *Am J Epidemiol*, 1995; **142**:909–917.

[393] Walker M, Shaper AG, Phillips AN, and Cook DG. Short stature, lung function and risk of a heart attack. *Int J Epidemiol*, 1989; **18**:602–606.

[394] Yarnell JW, Limb ES, Layzell JM, and Baker IA. Height: a risk marker for ischaemic heart disease: prospective results from the Caerphilly and Speedwell Heart Disease Studies. *Eur Heart J*, 1992; **13**:1602–1605.

[395] Kannam JP, Levy D, Larson M, and Wilson PW. Short stature and risk for mortality and cardiovascular disease events. The Framingham Heart Study. *Circulation*, 1994; **90**:2241–2247.

[396] Marenberg ME, Risch N, Berkman LF, Floderus B, and de Faire U. Genetic susceptibility to death from coronary heart disease in a study of twins. *N Engl J Med*, 1994; **330**:1041–1046.

[397] Ramser J, Seranski P, Hoff C, Poustka A, Reinhardt R, and Lehrach H. A physical map of the human genome. *Nature*, 2001; **409**:934–941.

[398] Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, and Sutton GG. The Sequence of the Human Genome. *Science*, 2001; **291** :1304–1351.

[399] Yano K, Reed DM, and McGee DL. Ten–year incidence of coronary heart disease in the Honolulu Heart Program. Relationship to biologic and lifestyle characteristics. *Am J Epidemiol*, 1984; **119**:653–666.

[400] National Center for Biotechnology Information. OMIM: Online mendelian inheritance in man. www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM (Professional genetic information about disease genes).

[401] Gu DF, Hinks LJ, Morton NE, and Day IN. The use of long PCR to confirm three common alleles at the CYP2A6 locus and the relationship between genotype and smoking habit. *Ann Hum Genet*, 2000; **64**:383–390.

[402] Humphries SE, Talmud PJ, Hawe E, Bolla M, Day INM, and Miller GJ. Apolipoprotein E4 and coronary heart disease in middle–aged men who smoke: a prospective study. *Lancet*, 2001; **358**:115–119.

[403] Davey Smith G and Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*, 2003; **32**:1–22.

[404] Davey Smith G, Harbord R, and Ebrahim S. Fibrinogen, C–reactive protein and coronary heart disease: does Mendelian randomization suggest the associations are non–causal? *QJM*, 2004; **97**:163–166.

[405] Little J and Khoury MJ. Mendelian randomisation: a new spin or real progress? *Lancet*, 2003; **362**:930–931.

[406] The Pooling Project Research Group. Relationship of blood pressure, serum cholesterol, smoking habit, relative weight and ECG abnormalities to incidence of major coronary events: final report of the pooling project. . *J Chron Dis*, 1978; **31**:201–306.

[407] Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, and Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*, 1998; **97**:1837–1847.

[408] Department of Health. National Service Framework for Coronary Heart Disease: modern standards and service models. London: The Stationery Office, 2000.

[409] Kannel WB, McGee D, and Gordon T. A general cardiovascular risk profile: the Framingham Study. *Am J Cardiol*, 1976; **38**:46–51.

[410] Anderson KM, Odell PM, Wilson PW, and Kannel WB. Cardiovascular disease risk profiles. *Am Heart J*, 1991; **121**:293–298.

[411] Shaper AG, Pocock SJ, Phillips AN, and Walker M. Identifying men at high risk of heart attacks: strategy for use in general practice. *BMJ*, 1986; **293**:474–479.

[412] Assman G. Lipid metabolism disorders and coronary heart disease. Munchen: MMV–Medizin–Verl, 1993.

[413] Tunstall-Pedoe H. The Dundee coronary risk–disk for management of change in risk factors. *BMJ*, 1991; **303**:744–747.

[414] Expert Panel on Detection Evaluation and Treatment of High Blood Cholesterol in Adults. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *JAMA*, 2001; **285**:2486–2497.

[415] Grundy SM, Pasternak R, Greenland P, Smith S Jr, and Fuster V. Assessment of cardiovascular risk by use of multiple–risk–factor assessment equations: a statement for healthcare professionals from the American Heart Association and the American College of Cardiology. *Circulation*, 1999; **100**:1481–1492.

[416] Fodor JG, Frohlich JJ, Genest JJ Jr, and McPherson PR. Recommendations for the management and treatment of dyslipidemia. Report of the Working Group on Hypercholesterolemia and Other Dyslipidemias. *CMAJ*, 2000; **162**:1441–1447.

[417] De Backer G, Ambrosioni E, Borch-Johnsen K, Brotons C, Cifkova R, Dallongeville J, Ebrahim S, Faergeman O, Graham I, and Mancia G. European guidelines on cardiovascular disease prevention in clinical practice: Third Joint Task Force of European and other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of eight societies and by invited experts). *Eur Heart J*, 2003; **24**:1601–1610.

[418] Conroy RM, Pyorala K, Fitzgerald AP, Sans S, Menotti A, De Backer G, De Bacquer D, Ducimetiere P, Jousilahti P, and Keil U. Estimation of ten–year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J*, 2003; **24**:987–1003.

[419] Menotti A, Puddu PE, and Lanti M. Comparison of the Framingham risk function–based coronary chart with risk function from an Italian population study. *Eur Heart J*, 2000; **21**:365–370.

[420] Laurier D, Nguyen PC, Cazelles B, and Segond P. Estimation of CHD risk in a French working population using a modified Framingham model. The PCV–METRA Group. *J Clinical Epidemiol*, 1994; **47**:1353–1364.

[421] Hense HW, Schulte H, Lowel H, Assmann G, and Keil U. Framingham risk function overestimates risk of coronary heart disease in men and women from Germany–results from the MONICA Augsburg and the PROCAM cohorts. *Eur Heart J*, 2003; **24**:937–945.

[422] Thomsen TF, McGee D, Davidsen M, and Jorgensen T. A cross–validation of risk–scores for coronary heart disease mortality based on data from the Glostrup Population Studies and Framingham Heart Study. *Int J Epidemiol*, 2002; **31**:817–822.

[423] Brindle P, Emberson J, Lampe F, Walker M, Whincup P, Fahey T, and Ebrahim S. Predictive accuracy of the Framingham coronary risk score in British men: prospective cohort study. *BMJ*, 2003; **327**:1267–1270.

[424] Empana JP, Ducimetiere P, Arveiler D, Ferrieres J, Evans A, Ruidavets JB, Haas B, Yarnell J, Bingham A, and Amouyel P. Are the Framingham and PROCAM coronary heart disease risk functions applicable to different European populations?: The PRIME Study. *Eur Ht J*, 2003; **24**:1903–1911.

[425] Wald NJ and Law MR. A strategy to reduce cardiovascular disease by more than 80%. *BMJ*, 2003; **326**:1419–1423.

[426] Vartiainen E, Jousilahti P, Alfthan G, Sundvall J, Pietinen P, and Puska P. Cardiovascular risk factor changes in Finland, 1972–1997. *Int J Epidemiol*, 2000; **29**:49–56.

[427] Dowse GK, Gareeboo H, Alberti KG, Zimmet P, Tuomilehto J, Purran A, Fareed D, Chitson P, and Collins VR. Changes in population cholesterol concentrations and other cardiovascular risk factor levels after five years of the non–communicable disease intervention programme in Mauritius. Mauritius Non–communicable Disease Study Group. *BMJ*, 1995; **311**:1255–1259.

[428] Ebrahim S and Davey Smith G. Lowering blood pressure: a systematic review of sustained effects of non–pharmacological interventions. *J Public Health Med*, 1998; **20**:441–448.

[429] Tang JL, Armitage JM, Lancaster T, Silagy CA, Fowler GH, Neil HAW, Smith George Davey, and Ebrahim Shah. Systematic review of dietary intervention trials to lower blood total cholesterol in free–living subjects * Commentary: Dietary change, cholesterol reduction, and the public health–what does meta–analysis add? *BMJ*, 1998; **316**:1213–1220.

[430] Forte JG, Miguel JM, Miguel MJ, de Padua F, and Rose G. Salt and blood pressure: a community trial. *J Hum Hypertens*, 1989; **3**:179–184.

[431] Ebrahim S and Davey Smith G. Exporting failure? Coronary heart disease and stroke in developing countries. *Int J Epidemiol*, 2001; **30**:201–205.

[432] Scientific Advisory Committee on Nutrition, Salt and Health. The Stationery Office, 2003. Available at http://www.sacn.gov.uk/pdfs/sacn_salt_final.pdf. Accessed December 10, 2004.

[433] Whelton PK, He J, Appel LJ, Cutler JA, Havas S, Kotchen TA, Roccella RS, Vallbona C, Winston MC, and Karimbakos J. National High Blood Pressure Education Program Coordinating Committee. Primary prevention of hypertension: clinical and public health advisory from The National High Blood Pressure Education Program. *JAMA*, 2002; **288**:1882–1888.

[434] He Feng J and MacGregor GA. How Far Should Salt Intake Be Reduced? *Hypertension*, 2003; **42**:1093–1099.

[435] Murray CJL, Lauer JA, Hutubessy RCW, Niessen L, Tomijima N, Rodgers A, Lawes CMM, and Evans DB. Effectiveness and costs of interventions to lower systolic blood pressure and cholesterol: a global and regional analysis on reduction of cardiovascular–disease risk. *Lancet*, 2003; **361**:717–725.

[436] Rodgers A and Neal B. Less salt does not necessarily mean less taste. *Lancet*, 1999; **353**:1332–1332.

[437] Weber M. Essays in Sociology. (Girth H, Mills CW, eds and translators). New York: Oxford University Press, 1946.

[438] Marx K. Capital. New York: International Publishers, 1967.

[439] Susser M, Watson W, and Hopper K. Sociology in Medicine. 3rd ed. New York: Oxford University Press, 1985.

[440] Liberatos P, Link BG, and Kelsey JL. The measurement of social class in epidemiology. *Epidemiol Rev*, 1988; **10**:87–121.

[441] Haan MN, Kaplan GA, and Syme SL. Socioeconomic status and health: old observations and new thoughts. In: Bunker JP, Gomby DS, Kehrer BM, eds. Pathways to Health: The role of social factors. Menlo Park, Calif: Henry J Kaiser Family Foundation, 1989.

[442] Classification of Occupations 1970. London: HM Stationery Office, 1970.

[443] Rose G and Marmot MG. Social class and coronary heart disease. *Br Heart J*, 1981; **45**:13–19.

[444] Liu K, Cedres LB, Stamler J, Dyer A, Stamler R, Nanas S, Berkson DM, Paul O, Lepper M, and Lindberg et al. Relationship of education to major risk factors and death from coronary heart disease, cardiovascular diseases and all causes, Findings of three Chicago epidemiologic studies. *Circulation*, 1982; **66**:1308–1314.

[445] Morris JN. Social inequalities undiminished. *Lancet*, 1979; **1**:87–90.

[446] Office for National Statistics. Mortality statistics: cause 1996, series DH2, no 23. London: The Stationery Office, 1998.

[447] Drever F and Bunting J. Patterns and trends in male mortality. In: Drever F, Whitehead M, eds. Health inequalities: decennial supplement: DS Series no 15. London: The Stationery Office, 1997.

[448] Pocock SJ, Shaper AG, Cook DG, Phillips AN, and Walker M. Social class differences in ischaemic heart disease in British men. *Lancet*, 1987; **2**:197–201.

[449] Diez-Roux AV, Nieto FJ, Tyroler HA, Crum LD, and Szklo M. Social inequalities and atherosclerosis. The atherosclerosis risk in communities study. *Am J Epidemiol*, 1995; **141**:960–972.

[450] van Rossum CTM, Shipley MJ, van de Mheen H, Grobbee DE, and Marmot MG. Employment grade differences in cause specific mortality. A 25 year follow up of civil servants from the first Whitehall study. *J Epidemiol Comm Health*, 2000; **54**:178–184.

[451] Woodward M, Shewry MC, Smith WC, and Tunstall-Pedoe H. Social status and coronary heart disease: results from the Scottish Heart Health Study. *Prev Med*, 1992; **21**:136–148.

[452] Gliksman MD, Kawachi I, Hunter D, Colditz GA, Manson JE, Stampfer MJ, Speizer FE, Willett WC, and Hennekens CH. Childhood socioeconomic status and risk of cardiovascular disease in middle aged US women: a prospective study. *J Epidemiol Comm Health*, 1995; **49**:10–15.

[453] Davey Smith G, Hart C, Blane D, Gillis C, and Hawthorne V. Lifetime socioeconomic position and mortality: prospective observational study. *BMJ*, 1997; **314**:547–552.

[454] Davey Smith G, McCarron P, Okasha M, and McEwen J. Social circumstances in childhood and cardiovascular disease mortality: prospective observational study of Glasgow University students. *J Epidemiol Comm Health*, 2001; **55**:340–.

[455] Kaplan GA and Salonen JT. Socioeconomic conditions in childhood and ischaemic heart disease during middle age. *BMJ*, 1990; **301**:1121–1123.

[456] Lynch JW, Kaplan GA, Cohen RD, Kauhanen J, Wilson TW, Smith NL, and Salonen JT. Childhood and adult socioeconomic status as predictors of mortality in Finland. *Lancet*, 1994; **343**:524–527.

[457] 1997–1999 World Health Statistics Annual. Geneva, Switzerland: World Health Organisation; 2000. Available at http://www-nt.who.int /whosis/statistics/menu.cfm?path=statistics, whsa&language = english. Accessed June 17, 2003.

[458] Criqui MH and Ringel BL. Does diet or alcohol explain the French paradox? *Lancet*, 1994; **344**:1719–1723.

[459] Law M and Wald N. Why heart disease mortality is low in France: the time lag explanation. *BMJ)*, 1999; **318**:1471–1476.

[460] Shimamoto T, Komachi Y, Inada H, Doi M, Iso H, Sato S, Kitamura A, Iida M, Konishi M, and Nakanishi et al. Trends for coronary heart disease and stroke and their risk factors in Japan. *Circulation*, 1989; **79**:503–515.

[461] Morris RW, Whincup PH, Lampe FC, Walker M, Wannamethee SG, and Shaper AG. Geographic variation in incidence of coronary heart disease in Britain: the contribution of established risk factors. *Heart*, 2001; **86**:277–283.

[462] Friedman GD. Cigarette smoking and geographic variation in coronary heart disease mortality in the United States. *J Chron Diseases*, 1967; **20**:769–779.

[463] Gupta R and Gupta VP. Meta–analysis of coronary heart disease prevalence in India. *Ind Heart J*, 1996; **48**:241–245.

[464] Tao SC, Huang ZD, Wu XG, Zhou BF, Xiao ZK, Hao JS, Li YH, Cen RC, and Rao XX. CHD and its risk factors in the People's Republic of China. *Int J Epidemiol*, 1989; **18**:S159–S163.

[465] Taylor R, Chey T, Bauman A, and Webster I. Socio–economic, migrant and geographic differentials in coronary heart disease occurrence in New South Wales. *Aus NZ J Pub Health*, 1999; **23**:20–26.

[466] Morris RW, Whincup PH, Emberson JR, Lampe FC, Walker M, and Shaper AG. North–South Gradients in Britain for Stroke and CHD: Are They Explained by the Same Factors? *Stroke*, 2003; **34**:2604–2609.

[467] He J, Klag MJ, Wu Z, and Whelton PK. Stroke in the People's Republic of China. I. Geographic variations in incidence and risk factors. *Stroke*, 1995; **26**:2222–2227.

[468] Stolley PD, Kuller LH, Nefzger MD, Tonascia S, Lilienfeld AM, Miller GD, and Diamond EL. Three–area epidemiological study of geographic differences in stroke mortality. II. Results. *Stroke*, 1977; **8**:551–557.

[469] Levin ML. The occurrence of lung cancer in man. *Acta Unio International Contra Cancrum*, 1953; **19**:531–541.

[470] Neter J, Kutner C, Nachtscheim C, and Wasserman W. Applied linear statistical models, Fourth edition. Irwin Press, 1995.

[471] McCullough P and Nelder JA. Generalized Linear Models, Second edition. Chapman and Hall, 1995.

[472] Cox DR and Oakes D. Analysis of survival data. London: Chapman and Hall, 1984.

[473] Efron B and Tibshirani RJ. An Introduction to the Bootstrap. London: Chapman and Hall, 1993.

[474] Kaplan EL and Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Ass*, 1958; **53**:457–481.

[475] Grambsch PM and Therneau TM. Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. *Biometrika*, 1994; **81**:515–526.

[476] Anderson P and Gill R. Cox's regression model for counting processes, a large sample study. *Ann Stat*, 1982; **10**:1100–1120.

[477] Easton DF, Peto J, and Babiker AG. Floating absolute risk: an alternative to relative risk in survival and case–control analysis avoiding an arbitrary reference group. *Stats Med*, 1991; **10** :1025–1035.

[478] Peto R. Birthweight as a risk factor for breast cancer (letter). *Lancet*, 1997; **349**:501.

[479] Michels KB, Robins JM, and Willett WC. Author's reply (Re: Birthweight as a risk factor for breast cancer). *Lancet*, 1997; **349**:502.

[480] Greenland S, Michels KB, Robins JM, Poole C, and Willett WC. Presenting statistical uncertainty in trends and dose–response relations. *Am J Epidemiol*, 1999; **149**:1077–1086.

[481] Easton D and Peto J. Re: Presenting statistical uncertainty in trends and dose–response relations (letter). *Am J Epidemiol*, 2000; **152**:393.

[482] Greenland S, Michels KB, Poole C, and Willett WC. Re: Presenting statistical uncertainty in trends and dose–response relations (letter). *Am J Epidemiol*, 2000; **152**:394.

[483] Schmitz PM, Habbema HDF, and Hermans J. The performance of logistic discrimination on myocardial infarction data, in comparison with some other discriminant analysis methods. *Stats Med*, 1983; **2**:199–205.

[484] Lachenbruch P. An almost unbiased method of obtaining confidence intervals for the probability opf missclassification in discriminant analysis. *Biometrics*, 1967; **23**:639–645.

[485] Efron B. The Jack–knife, the Bootstrap and other Re–sampling plans. Philadelphia: Society for Industrial and Applied Mathematics, 1982.

[486] Walker HM and Lev J. Statistical Inference. New York: Holt, Rinehart & Winston, 1953.

[487] Malinvaud E. Statistical Methods of Econometrics. Amsterdam: North–Holland Publishing Co., 1966.

[488] Cochran WG and William G. Errors of Measurement in Statistics. *Technometrics*, 1968; **10**:637–666.

[489] Kuha J. Corrections for exposure measurement error in logistic regression models with an application to nutritional data. *Stats Med*, 1994; **13**:1135–1148.

[490] Prentice RL. Covariate measurement errors and parameter estimation in a failure time regression model. *Biometrika*, 1982; **69**:331–342.

[491] Hughes MD. Regression dilution in the proportional hazards model. *Biometrics*, 1993; **49**:1056–1066.

[492] Fisher RA. Statistical Methods for Research Workers, 10th edn. Edinburgh: Oliver and Boyd, 1948.

[493] Duffy SW, Maximovitch DM, and Day NE. External validation, repeat determination, and precision of risk estimation in misclassified exposure data in epidemiology. *J Epidemiol Comm Health*, 1992; **46**:620–624.

[494] Pocock SJ, Shaper AG, Cook DG, Packham RF, Lacey RF, Powell P, and Russell PF. British Regional Heart Study: geographic variations in cardiovascular mortality, and the role of water quality. *BMJ*, 1980; **280**:1243–1249.

[495] Sigfusson N, Sigvaldason H, Steingrimsdottir L, Gudmundsdottir II, Stefansdottir I, Thorsteinsson T, and Sigurdsson G. Decline in ischaemic heart disease in Iceland and change in risk factor levels. *BMJ*, 1991; **302**:1371–1375.

[496] Elford J, Phillips AN, Thomson AG, and Shaper AG. Migration and geographic variations in ischaemic heart disease in Great Britain. *Lancet*, 1989; **1**:343–346.

[497] Bruce NG, Shaper AG, Walker M, and Wannamethee G. Observer bias in blood pressure studies. *J Hyperten*, 1988; **6**:375–380.

[498] MacFarlane PW and Lawrie TDV. An introduction to automated electrocardiogram interpretation. London: Butterworths, 1974.

[499] Whincup PH, Wannamethee G, MacFarlane PW, Walker M, and Shaper AG. Resting electrocardiogram and risk of coronary heart disease in middle–aged British men. *J Cardiovasc Risk*, 1995; **2**:533–543.

[500] MacFarlane PW, Lorimer AR, and Lawrie TD. 3 and 12 lead electrocardiogram interpretation by computer. A comparison on 1093 patients. *Br Ht J*, 1971; **33**:266–274.

[501] Thelle DS, Shaper AG, Whitehead TP, Bullock DG, Ashby D, and Patel I. Blood lipids in middle–aged British men. *Br Heart J*, 1983; **49**:205–213.

[502] Pocock SJ, Ashby D, Shaper AG, Walker M, and Broughton PM. Diurnal variations in serum biochemical and haematological measurements. *J Clin Pathol*, 1989; **42**:172–179.

[503] Andersen L, Dinesen B, Jorgensen PN, Poulsen F, and Roder ME. Enzyme immunoassay for intact human insulin in serum or plasma. *Clin Chem*, 1993; **39**:578–582.

[504] Feyerabend C and Russell MA. A rapid gas–liquid chromatographic method for the determination of cotinine and nicotine in biological fluids. *J Pharmacy and Pharmacology*, 1990; **42**:450–452.

[505] Danesh J, Whincup P, Walker M, Lennon L, Thomson A, Appleby P, Gallimore JR, and Pepys MB. Low grade inflammation and coronary heart disease: prospective study and updated meta–analyses. *BMJ*, 2000; **321**:199–204.

[506] Whincup PH, Refsum H, Perry IJ, Morris R, Walker M, Lennon L, Thomson A, Ueland PM, and Ebrahim SB. Serum total homocysteine and coronary heart disease: prospective study in middle aged men. *Heart*, 1999; **82**:448–454.

[507] Lampe FC, Whincup PH, Wannamethee SG, Ebrahim S, Walker M, and Shaper AG. Chest pain on questionnaire and prediction of major ischaemic heart disease events in men. *Eur Heart J*, 1998; **19**:63–73.

[508] Lampe FC. Chest pain on questionnaire and coronary heart disease in British men. PhD thesis. University of London, 2003.

[509] Cook DG, Shaper AG, and MacFarlane PW. Using the WHO (Rose) angina questionnaire in cardiovascular epidemiology. *Int J Epidemiol*, 1989; **18**:607-613.

[510] Ebrahim S, Papacosta O, Whincup P, Wannamethee G, Walker M, Nicolaides AN, Dhanjil S, Griffin M, Belcaro G, and Rumley et al. Carotid plaque, intima media thickness, cardiovascular risk factors, and prevalent cardiovascular disease in men and women: the British Regional Heart Study. *Stroke*, 1999; **30**:841–850.

[511] Walker M, Shaper AG, Lennon L, and Whincup PH. Twenty year follow–up of a cohort based in general practices in 24 British towns. *J Public Health Med*, 2000; **22**:479–485.

[512] Whincup PH, Bruce NG, Cook DG, and Shaper AG. The Dinamap 1846SX automated blood pressure recorder: comparison with the Hawksley random zero sphygmomanometer under field conditions. *J Epidemiol Comm Health*, 1992; **46**:164–169.

[513] Siedel J, Hagele EO, Ziegenhorn J, and Wahlefeld AW. Reagent for the enzymatic determination of serum total cholesterol with improved lipolytic efficiency. *Clin Chem*, 1983; **29**:1075–1080.

[514] Sugiuchi H, Uji Y, Okabe H, Uekama K, Kahahara N, and Miyauchi K. Direct measurement of high–density lipoprotein cholesterol in serum with polyethylene glycol modified enzymes and sulphated alpha–cyclodextrin. *Clin Chem*, 1995; **41**:717–723.

[515] Trinder P. Determination of blood glucose using 4–amino phenazone as oxygen acceptor. *J Clin Pathol*, 1969; **22**:246–246.

[516] Emberson JR, Whincup PH, Walker M, Thomas M, and Alberti KGMM. Biochemical measures in a population–based study: effect of fasting duration and time of day. *Ann Clin Biochem*, 2002; **39**:493–501.

[517] Eastern Stroke and Coronary Heart Disease Collaborative Research Group. Blood pressure, cholesterol, and stroke in eastern Asia. *Lancet*, 1998; **352**:1801–1807.

[518] Tanne D, Yaari S, and Goldbourt U. Risk profile and prediction of long–term ischemic stroke mortality: a 21–year follow–up in the Israeli Ischemic Heart Disease (IIHD) Project. *Circulation*, 1998; **98**:1365–1371.

[519] Salomaa V, Rasi V, Stengard J, Vahtera E, Pekkanen J, Vartiainen E, Ehnholm C, and Puska P. Intra– and interindividual variability of hemostatic factors and traditional cardiovascular risk factors in a three–year follow–up. *Thrombosis and Haemostasis*, 1998; **79**:969–974.

[520] Hart CL, Hole DJ, and Davey Smith G. Are two really better than one? Empirical examination of repeat blood pressure measurements and stroke risk in the Renfrew/Paisley and collaborative studies. *Stroke*, 2001; **32**:2697–2699.

[521] Iribarren C, Sharp D, Burchfiel CM, Sun P, and Dwyer JH. Association of serum total cholesterol with coronary disease and all–cause mortality: multivariate correction for bias due to measurement error. *Am J Epidemiol*, 1996; **143**:463–471.

[522] Lewington S, Thomsen T, Davidsen M, Sherliker P, and Clarke R. Regression dilution bias in blood total and high–density lipoprotein cholesterol and blood pressure in the Glostrup and Framingham prospective studies. *J Cardiovasc Risk*, 2003; **10**:143–148.

[523] Perry IJ, Wannamethee SG, Whincup PH, Shaper AG, Walker MK, and Alberti KG. Serum insulin and incident coronary heart disease in middle–aged British men. *Am J Epidemiol*, 1996; **144**:224–234.

[524] Farrell B, Godwin J, Richards S, and Warlow C. The United Kingdom transient ischaemic attack (UK–TIA) aspirin trial: final results. *J Neurol Neurosurg Psychiatry*, 1991; **54**:1044–1054.

[525] A comparison of two doses of aspirin (30 mg vs. 283 mg a day) in patients after a transient ischemic attack or minor ischemic stroke. The Dutch TIA Trial Study Group. *N Engl J Med*, 1991; **325**:1261–1266.

[526] Carotid Surgery Trialists' Collaborative Group. Randomised trial of endarterectomy for recently symptomatic carotid stenosis: final results of the MRC European Carotid Surgery Trial (ECST). *Lancet*, 1998; **351**:1379–1387.

[527] Howard SC and Rothwell PM. Regression dilution of systolic and diastolic blood pressure in patients with established cerebrovascular disease. *J Clin Epidemiol*, 2003; **56**:1084–1091.

[528] Clarke R, Lewington S, Donald A, Johnston C, Refsum H, Stratton I, Jacques P, Breteler MM, and Holman R. Underestimation of the importance of homocysteine as a risk factor for cardiovascular disease in epidemiological studies. *J Cardiovasc Risk*, 2001; **8**:363–369.

[529] Ridker PM, Rifai N, Pfeffer MA, Sacks F, and Braunwald E. Long–term effects of pravastatin on plasma concentration of C–reactive protein. The Cholesterol and Recurrent Events (CARE) Investigators. *Circulation*, 1999; **100**:230–235.

[530] Davey Smith G and Phillips AN. Inflation in epidemiology: 'The proof and measurement of association between two things' revisited. *BMJ*, 1996; **312**:1659–1661.

[531] Intersalt Co operative Research Group. Intersalt: an international study of electrolyte excretion and blood pressure. Results for 24 hour urinary sodium and potassium excretion. *BMJ*, 1988; **297**:319–328.

[532] Dyer AR, Elliott P, Marmot M, Kesteloot H, Stamler R, and Stamler J. Commentary: strength and importance of the relation of dietary salt to blood pressure. Intersalt Steering and Editorial Committee. *BMJ*, 1996; **312**:1661–1664.

[533] Clarke R, Lewington S, Youngman L, Sherliker P, Peto R, and Collins R. Underestimation of the importance of blood pressure and cholesterol for coronary heart disease mortality in old age. *Eur Heart J*, 2002; **23**:286–293.

[534] Prospective Studies Collaboration. Cholesterol, diastolic blood pressure, and stroke: 13,000 strokes in 450,000 people in 45 prospective cohorts. *Lancet*, 1995; **346**:1647–1653.

[535] Law MR, Wald NJ, Wu T, Hackshaw A, and Bailey A. Systematic underestimation of association between serum cholesterol concentration and ischaemic heart disease in observational studies: data from the BUPA study. *BMJ*, 1994; **308**:363–366.

[536] Hu FB, Stampfer MJ, Colditz GA, Ascherio A, Rexrode KM, Willett WC, and Manson JE. Physical activity and risk of stroke in women. *JAMA*, 2000; **283**:2961–2967.

[537] Jarvis MJ, Feyerabend C, Bryant A, Hedges B, and Primatesta P. Passive smoking in the home: plasma cotinine concentrations in non–smokers with smoking partners. *Tobacco Control*, 2001; **10**:368–374.

[538] Dunn G. Design and analysis of reliability studies: the statistical evaluation of measurement errors. London: Edward Arnold, 1989.

[539] Rose G. Strategy of prevention: lessons from cardiovascular disease. *BMJ*, 1981; **282**:1847–1851.

[540] Gueyffier F, Boutitie F, Boissel JP, Pocock S, Coope J, Cutler J, Ekbom T, Fagard R, Friedman L, and Perry et al. Effect of antihypertensive drug treatment on cardiovascular outcomes in women and men. A meta–analysis of individual patient data from randomized, controlled trials. The INDANA Investigators. *Ann Int Med*, 1997; **126**:761–767.

[541] Mant J and Hicks N. Detecting differences in quality of care: the sensitivity of measures of process and outcome in treating acute myocardial infarction. *BMJ*, 1995; **311**:793–796.

[542] Liao Y, McGee DL, Cooper RS, and Sutkowski MB. How generalizable are coronary risk prediction models? Comparison of Framingham and two national cohorts. *Am Heart J*, 1999; **137**:837–845.

[543] Tomasson H. Risk scores from logistic regression: unbiased estimates of relative and attributable risk. *Stats Med*, 1995; **14**:1331–1339.

[544] Vartiainen E, Puska P, Pekkanen J, Tuomilehto J, and Jousilahti P. Changes in risk factors explain changes in mortality from ischaemic heart disease in Finland. *BMJ*, 1994; **309**:23–27.

[545] Boreham R, Erns B, Falaschetti E, Hirani V, and Primatesta P. Risk factors for cardiovascular disease. London: The Stationery Office, 1998.

[546] Hooper L, Bartlett C, Davey Smith G, and Ebrahim S. Systematic review of long term effects of advice to reduce dietary salt in adults. *BMJ*, 2002; **325**:628–628.

[547] McCarron P, Okasha M, McEwen J, and Davey Smith G. Changes in blood pressure among students attending Glasgow University between 1948 and 1968: analyses of cross sectional surveys. *BMJ*, 2001; **322**:885–889.

[548] Capewell S, Morrison CE, and McMurray JJ. Contribution of modern cardiovascular treatment and risk factor changes to the decline in coronary heart disease mortality in Scotland between 1975 and 1994. *Heart*, 1999; **81**:380–386.

[549] Mullen PD. Compliance becomes concordance. *BMJ*, 1997; **314**:691-692.

[550] Smith WC, Lee AJ, Crombie IK, and Tunstall-Pedoe H. Control of blood pressure in Scotland: the rule of halves. *BMJ*, 1990; **300**:981-983.

[551] Teo KK, Yusuf S, Pfeffer M, Torp-Pedersen C, Kober L, Hall A, Pogue J, Latini R, Collins R, and ACE Inhibitors Collaborative Group. Effects of long–term treatment with angiotensin–converting–enzyme inhibitors in the presence or absence of aspirin: a systematic review. *Lancet*, 2002; **360**:1037–1043.

[552] Heller RF, Williams H, and Sittampalam Y. Social class and ischaemic heart disease: use of the male:female ratio to identify possible occupational hazards. *J Epidemiol Comm Health*, 1984; **38**:198–202.

[553] Davey Smith G, Shipley MJ, and Rose G. Magnitude and causes of socioeconomic differentials in mortality: further evidence from the Whitehall Study. *J Epidemiol Comm Health*, 1990; **44**:265–270.

[554] Freedman LS, Graubard BI, and Schatzkin A. Statistical validation of intermediate endpoints for chronic diseases. *Stats Med*, 1992; **11**:167–178.

[555] Breslow NE and Day NE. Statistical methods in cancer research. Volume 1 – the analysis of case–control studies. Lyon: International Agency for Research on Cancer, 1980.

[556] Krieger N, Williams DR, and Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annu.Rev.Public Health*, 1997; **18**:341–.

[557] Chandola T, Bartley M, Wiggins R, and Schofield P. Social inequalities in health by individual and household measures of social position in a cohort of healthy people. *J Epidemiol Comm Health*, 2003; **57**:56–62.

[558] Smith GD, Hart C, Watt G, Hole D, and Hawthorne V. Individual social class, area–based deprivation, cardiovascular disease risk factors, and mortality: the Renfrew and Paisley Study. *J Epidemiol Comm Health*, 1998; **52**:399–405.

[559] Liu K, Cedres LB, Stamler J, Dyer A, Stamler R, Nanas S, Berkson DM, Paul O, Lepper M, and Lindberg et al. Relationship of education to major risk factors and death from coronary heart disease, cardiovascular diseases and all causes, Findings of three Chicago epidemiologic studies. *Circulation*, 1982; **66**:1308–1314.

[560] Woodward M, Oliphant J, Lowe G, and Tunstall-Pedoe H. Contribution of contemporaneous risk factors to social inequality in coronary heart disease and all causes mortality. *Prev Med*, 2003; **36**:561–568.

[561] Law MR and Morris JK. Why is mortality higher in poorer areas and in more northern areas of England and Wales. *J Epidemiol Comm Health*, 1998; **52**:344-352.

[562] Greenland S. Ecologic versus individual–level sources of bias in ecologic estimates of contextual health effects. *Int J Epidemiol*, 2001; **30**:1343–1350.

[563] Jousilahti P, Tuomilehto J, Vartiainen E, Eriksson J, and Puska P. Relation of adult height to cause–specific and total mortality: a prospective follow–up study of 31,199 middle–aged men and women in Finland. *Am J Epidemiol*, 2000; **151**:1112–1120.

[564] Jarvis MJ and Wardle J. Social patterning of individual health behaviours: the case of cigarette smoking. In: Marmot M, Wilkinson RG. Social determinants of health. Oxford: Oxford Univesity Press, 1999.

[565] Health Survey for England 1998. London: The Stationery Office, 1999.

[566] Bradford-Hill A. The environment and disease: association or causation. *Proc R Soc Med*, 1965; **58**:295–290.

[567] Carroll RJ, Ruppert D, and Stefanski LA. Measurement Error in Non–linear models. New York: Chapman and Hall, 1995.

[568] World Health Statistics Annual 1990. Geneva: World Health Organization, 1991.

[569] Hopkins PN and Williams RR. Identification and relative weight of cardiovascular risk factors. *Cardiology Clinics*, 1986; **4**:3–31.

[570] Sterne JAC and Davey Smith G. Sifting the evidence – what's wrong with significance tests? *BMJ*, 2001; **322**:226–231.

[571] Nieto FJ. Infections and atherosclerosis: new clues from an old hypothesis? *Am J Epidemiol*, 1998; **148**:937–948.

[572] Barker DJP. Mothers, babies and health in later life. Edinburgh: Churchill Livingstone, 1998.

[573] Eriksson JG, Forsen T, Tuomilehto J, Osmond C, and Barker DJ. Early growth and coronary heart disease in later life: longitudinal study. *BMJ*, 2001; **322**:949–953.

[574] Danesh J and Peto R. Risk factors for coronary heart disease and infection with Helicobacter pylori: meta–analysis of 18studies. *BMJ*, 1998; **316**:1130–1132.

[575] Law MR, Wald NJ, and Morris JK. The performance of blood pressure and other cardiovascular risk factors as screening tests for ischaemic heart disease and stroke. *J Med Screen*, 2004; **11**:3–7.

[576] Franklin BA, Kahn JK, Gordon NF, and Bonow RO. A cardioprotective 'polypill'? Independent and additive benefits of lifestyle modification. *Am J Cardiol*, 2004; **94**:162–166.

[577] Evans A, Tolonen H, Hense HW, Ferrario M, Sans S, Kuulasmaa K, and WHO MONICA Project. Trends in coronary risk factors in the WHO MONICA project. *Int J Epidemiol*, 2001; **30**:S35–S40.

[578] Tunstall-Pedoe H, Vanuzzo D, Hobbs M, Mahonen M, Cepaitis Z, Kuulasmaa K, and Keil U. Estimation of contribution of changes in coronary care to improving survival, event rates, and coronary heart disease mortality across the WHO MONICA Project populations. *Lancet*, 2000; **355**:688–700.

[579] Tunstall-Pedoe H, Kuulasmaa K, Mahonen M, Tolonen H, Ruokokoski E, and Amouyel P. Contribution of trends in survival and coronary–event rates to changes in coronary heart disease mortality: 10–year results from 37 WHO MONICA project populations. Monitoring trends and determinants in cardiovascular disease. *Lancet*, 1999; **353**:1547–1557.

[580] Capewell S, Beaglehole R, Seddon M, and McMurray J. Explanation for the decline in coronary heart disease mortality rates in Auckland, New Zealand, between 1982 and 1993. *Circulation*, 2000; **102**:1511–1516.

[581] Kuulasmaa K, Tunstall-Pedoe H, Dobson A, Fortmann S, Sans S, Tolonen H, Evans A, Ferrario M, and Tuomilehto J. Estimation of contribution of changes in classic risk factors to trends in coronary–event rates across the WHO MONICA Project populations. *Lancet*, 2000; **355**:675–687.

[582] Jousilahti P, Vartiainen E, Tuomilehto J, Pekkanen J, and Puska P. Effect of risk factors and changes in risk factors on coronary mortality in three cohorts of middle–aged people in eastern Finland. *Am J Epidemiol*, 1995; **141**:50–60.

[583] Critchley JA and Capewell S. Substantial potential for reductions in coronary heart disease mortality in the UK through changes in risk factor levels. *J Epidemiol Comm Health*, 2003; **57**:243–247.

[584] Bots ML and Grobbee DE. Decline of coronary heart disease mortality in The Netherlands from 1978 to 1985: contribution of medical care and changes over time in presence of major cardiovascular risk factors. *J Cardiovasc Risk*, 1996; **3**:271–276.

[585] Goldman L and Cook EF. The decline in ischemic heart disease mortality rates. An analysis of the comparative effects of medical interventions and changes in lifestyle. *Ann Int Med*, 1984; **101**:825–836.

[586] Hunink MG, Goldman L, Tosteson AN, Mittleman MA, Goldman PA, Williams LW, Tsevat J, and Weinstein MC. The recent decline in mortality from coronary heart disease, 1980–1990. The effect of secular trends in risk factors and treatment. *JAMA*, 1997; **277**:535–542.

[587] Menotti A, Blackburn H, Kromhout D, Nissinen A, Fidanza F, Giampaoli S, Buzina R, Mohacek I, Nedeljkovic S, and Aravanis et al. Changes in population cholesterol levels and coronary heart disease deaths in seven countries. *Eur Heart J*, 1997; **18**:566–571.

[588] Unal B, Critchley JA, and Capewell S. Explaining the decline in coronary heart diseas mortality in England and Wales between 1981 and 2000. *Circulation*, 2004; **109**:1101–1107.

[589] Walker M, Shaper AG, and Cook DG. Non–participation and mortality in a prospective study of cardiovascular disease. *J Epidemiol Comm Health*, 1987; **41**:295–299.

[590] Office for National Statistics. Health Inequalities. London: The Stationery Office, 1997.

[591] Department of Health. The government's expenditure plans 2002–03 to 2003–04: Departmental report 2002. London: The Stationery Office, 2002.

[592] National public health expenditure report 1999–00. Canberra: AIHW, 2002.